```
                   GenCore version 5.1.6
            Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          July 22, 2004, 15:11:01 ; Search time 56 Seconds
                                           (without alignments)
                                           427.709 Million cell updates/sec

Title:           US-09-598-982-21
Perfect score:   1393
Sequence:        1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters:      283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        PIR_78:*
                 1:  pir1:*
                 2:  pir2:*
                 3:  pir3:*
                 4:  pir4:*

         Pred. No. is the number of results predicted by chance to have a
         score greater than or equal to the score of the result being printed,
         and is derived by analysis of the total score distribution.

                                    SUMMARIES
                      %
Result              Query
  No.     Score    Match Length DB ID                          Description
 ---------------------------------------------------------------------------
    1      1368     98.2    275  2  B35863                     tryptase (EC 3.4.2
    2      1363     97.8    275  2  A35863                     tryptase (EC 3.4.2
    3      1339     96.1    275  2  C35863                     tryptase (EC 3.4.2
    4    1232.5     88.5    274  2  A45754                     tryptase (EC 3.4.2
    5      1091     78.3    276  2  A38654                     mast cell proteina
    6      1077     77.3    275  2  A32410                     tryptase (EC 3.4.2
    7      1075     77.2    270  2  S56160                     mast cell tryptase
    8      1054     75.7    273  2  A47246                     tryptase (EC 3.4.2
    9      1051     75.4    274  2  JC4171                     tryptase (EC 3.4.2
   10      1005     72.1    237  2  S68702                     tryptase (EC 3.4.2
   11       807     57.9    230  2  I48685                     mast cell proteina
   12     684.5     49.1    269  2  B32410                     mastocytoma protei
   13       512     36.8    343  1  A57014                     prostasin (EC 3.4.
   14       467     33.5    625  1  KFHU1                      coagulation factor
   15       451     32.4    638  1  KQHUP                      plasma kallikrein
   16       448     32.2    455  2  A61545                     plasmin (EC 3.4.21
   17       447     32.1    638  1  KQMSPL                     plasma kallikrein
   18     446.5     32.1    812  1  PLBO                       plasmin (EC 3.4.21
   19       445     31.9    638  1  KQRTPL                     plasma kallikrein
   20     444.5     31.9    271  1  ELRT2                      pancreatic elastas
   21       443     31.8    245  1  KYBOB                      chymotrypsin (EC 3
   22       441     31.7    460  2  B61545                     plasmin (EC 3.4.21
   23       437     31.4   1524  2  T30337                     polyprotein - Afri
   24     436.5     31.3    269  2  B26823                     pancreatic elastas
   25     430.5     30.9    271  2  A25528                     pancreatic elastas
   26       429     30.8    810  1  PLHU                       plasmin (EC 3.4.21
   27       428     30.7    810  2  B30848                     plasmin (EC 3.4.21
   28       427     30.7    263  2  A21195                     chymotrypsin (EC 3
   29       427     30.7    266  1  ELPG                       pancreatic elastas
   30       426     30.6    263  2  A31299                     chymotrypsin (EC 3
   31       422     30.3    263  1  KYRTB                      chymotrypsin (EC 3
   32     420.5     30.2    812  1  PLMS                       plasmin (EC 3.4.21
   33       420     30.2    786  1  A47547                     serine proteinease
   34       419     30.1    245  1  KYBOA                      chymotrypsin (EC 3
   35       419     30.1    558  2  JC5878                     plasma hyaluronan-
   36       419     30.1    790  1  PLPG                       plasmin (EC 3.4.21
   37     416.5     29.9    246  2  B25528                     trypsin (EC 3.4.21
   38     414.5     29.8    246  1  TRRT2                      trypsin (EC 3.4.21
   39     414.5     29.8    264  2  I38136                     chymotrypsin-like
   40     413.5     29.7    367  2  JE0104                     testicular serine
   41     412.5     29.6    269  2  C26823                     pancreatic elastas
   42       412     29.6    258  4  S70439                     pancreatic elastas
   43       412     29.6    267  4  A56615                     probable pancreati
   44       411     29.5   1420  2  A32869                     apolipoprotein(a)
   45     409.5     29.4    246  1  TRRT1                      trypsin (EC 3.4.21

                                   ALIGNMENTS

RESULT 1
B35863
tryptase (EC 3.4.21.59) II precursor - human
N;Alternate names: tryptase beta
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 21-Jul-2000
C;Accession: B35863; A37193; I59473
R;Vanderslice, P.; Ballinger, S.M.; Tam, E.K.; Goldstein, S.M.; Craik, C.S.; Caughey, G.H
Proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990
A;Title: Human mast cell tryptase: multiple cDNAs and genes reveal a multigene serine pro
A;Reference number: A35863; MUID:90251647; PMID:2187193
A;Accession: B35863
A;Molecule type: mRNA; DNA
A;Residues: 1-275 <VAN>
A;Cross-references: GB:M33492; NID:g339982; PIDN:AAA36779.1; PID:g339983
A;Note: residues 2-275 are derived from mRNA; residue one was inferred from the genomic I
R;Miller, J.S.; Moxley, G.; Schwartz, L.B.
J. Clin. Invest. 86, 864-870, 1990
A;Title: Cloning and characterization of a second complementary DNA for human tryptase.
A;Reference number: A37193; MUID:90369005; PMID:2203827
A;Accession: A37193
A;Molecule type: mRNA
A;Residues: 1-275 <MIL>
A;Cross-references: GB:M37488; NID:g179583; PIDN:AAA51843.1; PID:g179584
R;Blom, T.; Hellman, L.
Scand. J. Immunol. 37, 203-208, 1993
A;Title: Characterization of a tryptase mRNA expressed in the human basophil cell line KU
A;Reference number: I59473; MUID:93166209; PMID:8434231
A;Accession: I59473
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-275 <RES>
A;Cross-references: GB:S55551; NID:g265666; PIDN:AAD13876.1; PID:g4261576
A;Experimental source: basophil cell line KU812
C;Genetics:
A;Gene: GDB:TPS1
A;Cross-references: GDB:125890; OMIM:191080
A;Map position: 16pter-16qter
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-30/Domain: activation peptide #status predicted <ACT>
F;31-275/Product: tryptase I #status predicted <MAT>
F;31-267/Domain: trypsin homology <TRY>
F;74,121,224/Active site: His, Asp, Ser #status predicted

  Query Match             98.2%;  Score 1368;  DB 2;  Length 275;
  Best Local Similarity   99.6%;  Pred. No. 1.1e-119;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
                ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
```

```
Db          31 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy          65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 150

Qy         125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy         185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy         245 VPKKP 249
               |||||
Db         271 VPKKP 275


RESULT 2
A35863
tryptase (EC 3.4.21.59) I precursor - human
C;Species: Homo sapiens (man)
C;Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 21-Jul-2000
C;Accession: A35863; D35863; A60939; A39326
R;Vanderslice, P.; Ballinger, S.M.; Tam, E.K.; Goldstein, S.M.; Craik, C.S.; Caughey, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990
A;Title: Human mast cell tryptase: multiple cDNAs and genes reveal a multigene serine pr
A;Reference number: A35863; MUID:90251647; PMID:2187193
A;Accession: A35863
A;Molecule type: DNA
A;Residues: 1-275 <VAN>
A;Cross-references: GB:M33494; NID:g3927804; PIDN:AAC83172.1; PID:g339977
A;Accession: D35863
A;Molecule type: mRNA
A;Residues: 1-275 <VA2>
A;Cross-references: GB:M33491
R;Butterfield, J.H.; Weiler, D.A.; Hunt, L.W.; Wynn, S.R.; Roche, P.C.
J. Leukoc. Biol. 47, 409-419, 1990
A;Title: Purification of tryptase from a human mast cell line.
A;Reference number: A60939; MUID:90244210; PMID:2110591
A;Accession: A60939
A;Molecule type: protein
A;Residues: 31-38,'P',40-41,'X',43,'T',45-48,'X',50 <BUT>
A;Experimental source: mast cell
A;Note: 44-Gly was also found
R;Cromlish, J.A.; Seidah, N.G.; Marcinkiewicz, M.; Hamelin, J.; Johnson, D.A.; Chretien,
J. Biol. Chem. 262, 1363-1373, 1987
A;Title: Human pituitary tryptase: molecular forms, NH-2-terminal sequence, immunocytoch
A;Reference number: A39326; MUID:87109258; PMID:3543004
A;Accession: A39326
A;Molecule type: protein
A;Residues: 31-38 <CRO>
A;Experimental source: pituitary
C;Genetics:
A;Introns: 21/1; 78/2; 177/1; 221/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-30/Domain: activation peptide #status predicted <ACT>
F;31-275/Product: tryptase I #status experimental <MAT>
F;31-267/Domain: trypsin homology <TRY>
F;74,121,224/Active site: His, Asp, Ser #status predicted

  Query Match             97.8%;  Score 1363;  DB 2;  Length 275;
  Best Local Similarity   99.2%;  Pred. No. 3.2e-119;
  Matches  243;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy           5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
               ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          31 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy          65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
```

```
               ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFP 150

Qy         125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy         185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy         245 VPKKP 249
               |||||
Db         271 VPKKP 275


RESULT 3
C35863
tryptase (EC 3.4.21.59) III precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-Mar-2003
C;Accession: C35863; E35863; A38893
R;Vanderslice, P.; Ballinger, S.M.; Tam, E.K.; Goldstein, S.M.; Craik, C.S.; Caughey, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 3811-3815, 1990
A;Title: Human mast cell tryptase: multiple cDNAs and genes reveal a multigene serine pr
A;Reference number: A35863; MUID:90251647; PMID:2187193
A;Accession: C35863
A;Molecule type: mRNA
A;Residues: 9-275 <VAN>
A;Accession: E35863
A;Molecule type: DNA
A;Residues: 1-9 <VA2>
A;Cross-references: GB:M33494; NID:g3927804; PIDN:AAC83172.1; PID:g339977
A;Note: the first nine residues of this sequence are inferred from genomic DNA of trypta
R;Vanderslice, P.
submitted to GenBank, April 1990
A;Reference number: A38893
A;Accession: A38893
A;Molecule type: mRNA
A;Residues: 9-131,'K',132-275 <VA3>
A;Cross-references: GB:M33493; NID:g339984; PIDN:AAA36780.1; PID:g339985
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-30/Domain: activation peptide #status predicted <ACT>
F;31-275/Product: tryptase I #status predicted <MAT>
F;31-267/Domain: trypsin homology <TRY>
F;74,121,224/Active site: His, Asp, Ser #status predicted

  Query Match             96.1%;  Score 1339;  DB 2;  Length 275;
  Best Local Similarity   98.0%;  Pred. No. 5.4e-117;
  Matches  240;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy           5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
               ||||||||||||||||||||   |||||||||||||||||||| ||||||||||||||||
Db          31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy          65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
               ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db          91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFP 150

Qy         125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy         185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy         245 VPKKP 249
               |||||
Db         271 VPKKP 275
```

```
RESULT 4
A45754
tryptase (EC 3.4.21.59) alpha precursor - human
C;Species: Homo sapiens (man)
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 08-Sep-1997
C;Accession: A45754; B37193
R;Miller, J.S.; Westin, E.H.; Schwartz, L.B.
J. Clin. Invest. 84, 1188-1195, 1989
A;Title: Cloning and characterization of complementary DNA for human tryptase.
A;Reference number: A45754; MUID:90009311; PMID:2677049
A;Accession: A45754
A;Molecule type: mRNA
A;Residues: 1-274 <MIL>
A;Cross-references: GB:M30038
R;Miller, J.S.; Moxley, G.; Schwartz, L.B.
J. Clin. Invest. 86, 864-870, 1990
A;Title: Cloning and characterization of a second complementary DNA for human tryptase.
A;Reference number: A37193; MUID:90369005; PMID:2203827
A;Accession: B37193
A;Molecule type: mRNA
A;Residues: 1-274 <MI2>
A;Cross-references: GB:M30038
A;Note: the sequence from Fig. 4 is inconsistent with that from Fig. 2 in having 205-Pr
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-30/Domain: activation peptide #status predicted <ACT>
F;31-274/Product: tryptase I #status predicted <MAT>
F;31-266/Domain: trypsin homology <TRY>
F;74,120,223/Active site: His, Asp, Ser #status predicted

  Query Match             88.5%;  Score 1232.5;  DB 2;  Length 274;
  Best Local Similarity   89.8%;  Pred. No. 4.4e-107;
  Matches  220;  Conservative    7;  Mismatches   17;  Indels    1;  Gaps    1;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| |:|||||||| |||
Db         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVN- 89

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
                 ||||||||||||||:||||||  | ||||||||||||| :|| |||| |||||||||
Db         90 SGTHLYYQDQLLPVSRIMVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 149

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db        150 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 209

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||::||||:||||||||||||||||||||||| |||||||||||||||||||||||||
Db        210 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 269

Qy        245 VPKKP 249
              |||||
Db        270 VPKKP 274


RESULT 5
A38654
mast cell proteinase 6 (EC 3.4.21.-) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Feb-1992 #sequence_revision 17-Feb-1994 #text_change 22-Jun-1999
C;Accession: A38654; B38654; D35646; I59478
R;Reynolds, D.S.; Gurley, D.S.; Austen, K.F.; Serafin, W.E.
J. Biol. Chem. 266, 3847-3853, 1991
A;Title: Cloning of the cDNA and gene of mouse mast cell protease-6. Transcription by pr
A;Reference number: A38654; MUID:91139682; PMID:1995638
A;Accession: A38654
A;Molecule type: DNA
A;Residues: 1-276 <REY>
A;Cross-references: GB:M57625; NID:g200506; PIDN:AAA39987.1; PID:g200507
A;Note: the authors translated the codon CGC for residue 24 as Ala, GAG for residue 37 a
s Gly, GAG for residue 148 as Gly, GAG for residue 168 as Gly, and GAA for 185 as Gly
A;Accession: B38654
A;Molecule type: mRNA
A;Residues: 1-276 <RE2>
A;Cross-references: GB:M57626; NID:g200508; PIDN:AAA39988.1; PID:g200509
R;Reynolds, D.S.; Stevens, R.L.; Lane, W.S.; Carr, M.H.; Austen, K.F.; Serafin, W.E.
Proc. Natl. Acad. Sci. U.S.A. 87, 3230-3234, 1990
A;Title: Different mouse mast cell populations express various combinations of at least
A;Reference number: A35646; MUID:90222202; PMID:2326280
A;Accession: D35646
A;Molecule type: protein
A;Residues: 32-54 <RE3>
R;Huang, R.; Abrink, M.; Gobl, A.E.; Nilsson, G.; Aveskogh, M.; Larsson, L.G.; Nilsson,
Scand. J. Immunol. 38, 359-367, 1993
A;Title: Expression of a mast cell tryptase in the human monocytic cell lines U-937 and
A;Reference number: I59478; MUID:94023807; PMID:8210998
A;Accession: I59478
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-276 <RES>
A;Cross-references: GB:L31853; NID:g473480; PIDN:AAA39725.1; PID:g473481
C;Genetics:
A;Gene: MMCP-6
A;Introns: 24/1; 79/2; 168/1; 222/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-31/Domain: activation peptide #status predicted <ACT>
F;32-276/Product: mast cell proteinase 6 #status experimental <MAT>
F;32-268/Domain: trypsin homology <TRY>
F;75,122,225/Active site: His, Asp, Ser #status predicted

  Query Match             78.3%;  Score 1091;  DB 2;  Length 276;
  Best Local Similarity   78.2%;  Pred. No. 6.6e-94;
  Matches  190;  Conservative   19;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||| ||  ||||||||||    ||:||||||||||||||||| |||| :|      ||||
Db         32 IVGGHEASESKWPWQVSLRFKLNYWIHFCGGSLIHPQWVLTAAHCVGPHIKSPQLFRVQL 91

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |||:||| |||| ::||:||| :|||: |||:|||||| || ||:|:| ::|||||||||
Db         92 REQYLYYGDQLLSLNRIVVHPHYYTAEGGADVALLELEVPVNVSTHIHPISLPPASETFP 151

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||  ||||||||:|||| ||||:|||||||||:|| :|| ||| | |||||  || | ||
Db        152 PGTSCWVTGWGDIDNDEPLPPPYPLKQVKVPIVENSLCDRKYHTGLYTGDDFPIVHDGML 211

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              |||||||||||||||||||||| |||||||||||||||||||:||||||||||||||| |
Db        212 CAGNTRRDSCQGDSGGPLVCKVKGTWLQAGVVSWGEGCAQPNKPGIYTRVTYYLDWIHRY 271

Qy        245 VPK 247
              ||:
Db        272 VPE 274


RESULT 6
A32410
tryptase (EC 3.4.21.59) precursor - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 12-Oct-1989 #sequence_revision 12-Oct-1989 #text_change 22-Jun-1999
C;Accession: A32410
R;Vanderslice, P.; Craik, C.S.; Nadel, J.A.; Caughey, G.H.
Biochemistry 28, 4148-4155, 1989
A;Title: Molecular cloning of dog mast cell tryptase and a related protease: structural
A;Reference number: A32410; MUID:89352460; PMID:2504277
A;Accession: A32410
A;Molecule type: mRNA
A;Residues: 1-275 <VAN>
A;Cross-references: GB:M24664; NID:g163982; PIDN:AAA30854.1; PID:g163983; GB:J02862
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-21/Domain: signal sequence #status predicted <SIG>
```

```
F;22-30/Domain: activation peptide #status predicted <ACT>
F;31-275/Product: tryptase #status predicted <MAT>
F;31-267/Domain: trypsin homology <TRY>
F;74,121,224/Active site: His, Asp, Ser #status predicted

  Query Match             77.3%;  Score 1077;  DB 2;  Length 275;
  Best Local Similarity   77.1%;  Pred. No. 1.3e-92;
  Matches  189;  Conservative   21;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||:||| ||||||||||: | || | ||||||||||||||| ||||:|     :||||
Db         31 IVGGREAPGSKWPWQVSLRLKGQYWRHICGGSLIHPQWVLTAAHCVGPNVVCPEEIRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||| ||||:||::|| :|| : ||||||||||:|| ||:||  |||||| :|||
Db         91 REQHLYYQDHLLPVNRIVMHPNYYTPENGADIALLELEDPVNVSAHVQPVTLPPALQTFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               | ||||||||||  :   ||||||||||||||:|| :|| :||||  ||| |||||:|||
Db        151 TGTPCWVTGWGDVHSGTPLPPPFPLKQVKVPIVENSMCDVQYHLGLSTGDGVRIVREDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||:: |||||||||||||:| | ||||||||||||||||||||||||| ||||||| |
Db        211 CAGNSKSDSCQGDSGGPLVCRVRGVWLQAGVVSWGEGCAQPNRPGIYTRVAYYLDWIHQY 270

Qy        245 VPKKP 249
              |||:|
Db        271 VPKEP 275


RESULT 7
S56160
mast cell tryptase precursor - Mongolian jird
C;Species: Meriones unguiculatus (Mongolian jird)
C;Date: 27-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 20-Jun-2000
C;Accession: S56160
R;Murakumo, Y.; Ide, H.; Itoh, H.; Tomita, M.; Kobayashi, T.; Maruyama, H.; Horii, Y.;
Biochem. J. 309, 921-926, 1995
A;Title: Cloning of the cDNA encoding mast cell tryptase of Mongolian gerbil, Meriones
A;Reference number: S56160; MUID:95366971; PMID:7639711
A;Accession: S56160
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-270 <MUR>
A;Cross-references: EMBL:D31789; NID:g517122; PIDN:BAA06598.1; PID:g517123
C;Superfamily: trypsin; trypsin homology
F;26-262/Domain: trypsin homology <TRY>

  Query Match             77.2%;  Score 1075;  DB 2;  Length 270;
  Best Local Similarity   77.4%;  Pred. No. 2e-92;
  Matches  188;  Conservative   18;  Mismatches   37;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||||||| :||||||||| :  || ||||||||||||||||| |||| : |   :||||
Db         26 IVGGQEAPGNKWPWQVSLRANETYWRHFCGGSLIHPQWVLTAAHCVGPTIADPNKVRVQL 85

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |:|:||| | || ||||| || ||  | |||||||||: || :||||| |:|||||||||
Db         86 RKQYLYYHDHLLAVSRIITHPTFYATQNGADIALLELKNPVNISSHVHPVSLPPASETFP 145

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               |  |||||||::|||  ||||||||:|:||::|| :|| ||| | ||||:: |||||||
Db        146 SGTLCWVTGWGNIDNDVSLPPPFPLKEVQVPVVENQLCDLKYHKGVYTGDNIHIVRDDML 205

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||   |||||||||||||||||||||||||||||||| |||||||||||||||||| |
Db        206 CAGNEGHDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCALPNRPGIYTRVTYYLDWIHRY 265

Qy        245 VPK 247
              |||
Db        266 VPK 268
```

```
RESULT 8
A47246
tryptase (EC 3.4.21.59) 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A47246
R;McNeil, H.P.; Reynolds, D.S.; Schiller, V.; Ghildyal, N.; Gurley, D.S.; Austen, K.F.;
Proc. Natl. Acad. Sci. U.S.A. 89, 11174-11178, 1992
A;Title: Isolation, characterization, and transcription of the gene encoding mouse mast
A;Reference number: A47246; MUID:93087489; PMID:1454796
A;Accession: A47246
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-273 <MCN>
A;Cross-references: GB:L00653; NID:g200518; PIDN:AAA39992.1; PID:g200519
A;Note: sequence extracted from NCBI backbone (NCBIN:119745, NCBIP:119746)
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;29-265/Domain: trypsin homology <TRY>

  Query Match             75.7%;  Score 1054;  DB 2;  Length 273;
  Best Local Similarity   76.5%;  Pred. No. 1.8e-90;
  Matches  186;  Conservative   17;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||||||  :||||||||| :  |||||||||||||||||||| |||||| |   :||||
Db         29 IVGGQEAHGNKWPWQVSLRANDTYWMHFCGGSLIHPQWVLTAAHCVGPDVADPNKVRVQL 88

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |:|:||| | |: ||:|| || ||  | |||||||:|  || :| :|| | |||||||||
Db         89 RKQYLYYHDHLMTVSQIITHPDFYIVQDGADIALLKLTNPVNISDYVHPVPLPPASETFP 148

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               |  |||||||::||   ||||||||:|:|||:|||:|| ||| |  |||:| |||||||
Db        149 SGTLCWVTGWGNIDNGVNLPPPFPLKEVQVPIIENHLCDLKYHKGLITGDNVHIVRDDML 208

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||   |||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        209 CAGNEGHDSCQGDSGGPLVCKVEDTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 268

Qy        245 VPK 247
              |||
Db        269 VPK 271


RESULT 9
JC4171
tryptase (EC 3.4.21.59) precursor - rat
N;Alternate names: mast cell tryptase
C;Species: Rattus norvegicus (Norway rat)
C;Date: 27-Aug-1995 #sequence_revision 27-Oct-1995 #text_change 20-Jun-2000
C;Accession: JC4171
R;Ide, H.; Itoh, H.; Tomita, M.; Murakumo, Y.; Kobayashi, T.; Maruyama, H.; Osada, Y.;
J. Biochem. 118, 210-215, 1995
A;Title: cDNA sequencing and expression of rat mast cell tryptase.
A;Reference number: JC4171; MUID:96015171; PMID:8537314
A;Accession: JC4171
A;Molecule type: mRNA
A;Residues: 1-274 <IDE>
A;Cross-references: DDBJ:D38455; NID:g556555; PIDN:BAA07486.1; PID:g556556
C;Comment: This enzyme is basically specific for a connective tissue mast cell, it is up-
einase inhibitors.
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; hydrolase; mast cell; serine proteinase; zymogen
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-29/Domain: activation peptide #status predicted <ACT>
F;30-274/Product: mast cell tryptase #status predicted <MAT>
F;30-266/Domain: trypsin homology <TRY>
F;73,120,223/Active site: His, Asp, Ser #status predicted
F;131/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             75.4%;  Score 1051;  DB 2;  Length 274;
  Best Local Similarity   75.8%;  Pred. No. 3.4e-90;
```

```
  Matches  185;  Conservative   21;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||:||  ||||||||||    :||||||||||||||||||| |||  :|     ||||
Db         30 IVGGREASESKWPWQVSLRFKFSFWMHFCGGSLIHPQWVLTAAHCVGLHIKSPELFRVQL 89

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |||:||| |||| |:| :||| :|| : |||||||||| || ||:|:| ::|||||||||
Db         90 REQYLYYADQLLTVNRTVVHPHYYTVEDGADIALLELEIPVNVSTHIHPISLPPASETFP 149

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               |  ||||||||:|:|| | ||:|||||||||:|| :|| ||| | |||||| ||:| ||
Db        150 SGTSCWVTGWGDIDSDEPLLPPYPLKQVKVPIVENSLCDRKYHTGLYTGDDVPIVQDGML 209

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              |||||| ||||||||||||||| ||||||||||||||||: ||||||||||||||||| |
Db        210 CAGNTRSDSCQGDSGGPLVCKVKGTWLQAGVVSWGEGCAEANRPGIYTRVTYYLDWIHRY 269

Qy        245 VPKK 248
              ||::
Db        270 VPQR 273


RESULT 10
S68702
tryptase (EC 3.4.21.59) - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 25-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 20-Jun-2000
C;Accession: S68702; S68676
R;Pallaoro, M.; Gambacurta, A.; Fiorucci, L.; Mignogna, G.; Barra, D.; Ascoli, F.
Eur. J. Biochem. 237, 100-105, 1996
A;Title: cDNA cloning and primary structure of tryptase from bovine mast cells, and evid
A;Reference number: S68676; MUID:96203914; PMID:8620861
A;Accession: S68702
A;Molecule type: mRNA
A;Residues: 1-237 <PAL>
A;Cross-references: EMBL:X94982; NID:g1332446; PIDN:CAA64438.1; PID:g1332447
A;Accession: S68676
A;Molecule type: protein
A;Residues: 'IVGGQEAP',1-8;61-74;90-97;126-148;162-190;208-222;227-235 <PAW>
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; hydrolase; serine proteinase; zymogen
F;1-229/Domain: trypsin homology (fragment) <TRY>
F;36,83,177/Active site: His, Asp, Ser #status predicted
F;94,195/Binding site: carbohydrate (Asn) (covalent) #status predicted

  Query Match             72.1%;  Score 1005;  DB 2;  Length 237;
  Best Local Similarity   75.4%;  Pred. No. 5.5e-86;
  Matches  178;  Conservative   21;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         14 SKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQLREQHLYYQD 73
              |:|||||||||   || | ||||||||||||||| ||||:|   :  |||||||||||||
Db          2 SQWPWQVSLRVSRRYWRHHCGGSLIHPQWVLTAAHCVGPEVHGPSYFRVQLREQHLYYQD 61

Qy         74 QLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTG 133
              ||||:|||| ||  |: : |||||||||:: | :| ||  ||||| |||||||  |||||
Db         62 QLLPISRIIPHPNCYSVKNGADIALLELDKLVNISWHVQPVTLPPESETFPPGTQCWVTG 121

Qy        134 WGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDS 193
              ||:|||  ||||||||||||||::|| :|| ||| |  |||:| |||:||||||:: |:
Db        122 WGNVDNGRRLPPPFPLKQVKVPVVENSVCDRKYHSGLSTGDNVPIVREDMLCAGDSGRNF 181

Qy        194 CQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 249
              ||||||||||||||||||||||||||:|||:||||||||||| |||||| |||: |
Db        182 CQGDSGGPLVCKVNGTWLQAGVVSWGDGCAKPNRPGIYTRVTSYLDWIHQYVPQGP 237


RESULT 11
I48685
mast cell proteinase 6 (EC 3.4.21.-) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 22-Jun-1999
C;Accession: I48685; S43172
```

```
R;Huang, R.; Hellman, L.
Immunogenetics 40, 397-414, 1994
A;Title: Genes for mast-cell serine protease and their molecular evolution.
A;Reference number: I48684; MUID:95048582; PMID:7959952
A;Accession: I48685
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-230 <RES>
A;Cross-references: EMBL:X78542; NID:g468809; PIDN:CAA55288.1; PID:g468810
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;32-230/Domain: trypsin homology #status atypical <TRY>

  Query Match             57.9%;  Score 807;  DB 2;  Length 230;
  Best Local Similarity   74.3%;  Pred. No. 1.4e-67;
  Matches  142;  Conservative   17;  Mismatches   32;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||| ||  ||||||||||    ||:||||||||||||||||| |||| :|     ||||
Db         32 IVGGHEASESKWPWQVSLRFKLNYWIHFCGGSLIHPQWVLTAAHCVGPHIKSPQLFRVQL 91

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |||:||| |||| ::||:||| :|||: |||:|||||| || ||:|:| ::|||||||||
Db         92 REQYLYYGDQLLSLNRIVVHPHYYTAEGGADVALLELEVPVNVSTHIHPISLPPASETFP 151

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||  ||||||||:|||| ||||:|||||||||:|| :|| ||| | |||||  || | ||
Db        152 PGTSCWVTGWGDIDNDEPLPPPYPLKQVKVPIVENSLCDRKYHTGLYTGDDFPIVHDGML 211

Qy        185 CAGNTRRDSCQ 195
              |||||||||||
Db        212 CAGNTRRDSCQ 222


RESULT 12
B32410
mastocytoma proteinase (EC 3.4.21.-) precursor - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 12-Oct-1989 #sequence_revision 31-Dec-1993 #text_change 10-Sep-1997
C;Accession: B32410
R;Vanderslice, P.; Craik, C.S.; Nadel, J.A.; Caughey, G.H.
Biochemistry 28, 4148-4155, 1989
A;Title: Molecular cloning of dog mast cell tryptase and a related protease: structural
A;Reference number: A32410; MUID:89352460; PMID:2504277
A;Accession: B32410
A;Molecule type: mRNA
A;Residues: 1-269 <VAN>
A;Cross-references: GB:M24665; NID:g163984; PID:g163985; GB:J02862
A;Note: the authors translated the codon ACG for residue 114 as Ser
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-269/Product: mastocytoma proteinase #status predicted <MAT>
F;20-259/Domain: trypsin homology <TRY>
F;66,116,217/Active site: His, Asp, Ser #status predicted

  Query Match             49.1%;  Score 684.5;  DB 2;  Length 269;
  Best Local Similarity   52.6%;  Pred. No. 4.4e-56;
  Matches  132;  Conservative   32;  Mismatches   78;  Indels    9;  Gaps    4;

Qy          5 IVGGQEAPRSKWPWQVSLRVHG---PYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALR 61
              |||| : |  ::||||||| ||     | | ||||||||||||||| ||  :  : |  ||
Db         20 IVGGCKVPARRYPWQVSLRFHGMGSGQWQHICGGSLIHPQWVLTAAHCVELEGLEAATLR 79

Qy         62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIG---ADIALLELEEPVKVSSHVHTVTLPP 118
              ||: :  ||  |||  |: || || |  :  |   ||||||:|| |: :|  |: |:||
Db         80 VQVGQLRLYDHDQLCNVTEIIRHPNFNMSWYGWDTADIALLKLEAPLTLSEDVNLVSLPS 139

Qy        119 ASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRI 178
               |   |||| ||||||||: :   ||||: |::|:|||: |  |:  |        |  :
Db        140 PSLIVPPGMLCWVTGWGDIADHTPLPPPYHLQEVEVPIVGNRECNCHYQ--TILEQDDEV 197

Qy        179 VRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYL 238
```

```
              :: ||||||:   |||| ||||||||:   ||:| |||||| ||   | ||:| ||| |:
Db        198 IKQDMLCAGSEGHDSCQMDSGGPLVCRWKCTWIQVGVVSWGYGCGY-NLPGVYARVTSYV 256

Qy        239 DWIHHYVPKKP 249
               ||| ::|  |
Db        257 SWIHQHIPLSP 267

RESULT 13
A57014
prostasin (EC 3.4.21.-) precursor - human
C;Species: Homo sapiens (man)
C;Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 21-Apr-2003
C;Accession: A57014; A54866
R;Yu, J.X.; Chao, L.; Chao, J.
J. Biol. Chem. 270, 13483-13489, 1995
A;Title: Molecular cloning, tissue-specific expression, and cellular localization of hum
A;Reference number: A57014; MUID:95286644; PMID:7768952
A;Accession: A57014
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-343 <RES>
A;Cross-references: GB:L41351; NID:g862304; PIDN:AAC41759.1; PID:g862305
A;Experimental source: prostate
A;Note: parts of this sequence were determined by protein sequencing
R;Yu, J.X.; Chao, L.; Chao, J.
J. Biol. Chem. 269, 18843-18848, 1994
A;Title: Prostasin is a novel human serine proteinase from seminal fluid. Purification,
A;Reference number: A54866; MUID:94308140; PMID:8034638
A;Accession: A54866
A;Molecule type: protein
A;Residues: 45-64 <YUA>
C;Genetics:
A;Gene: GDB:PRSS8
A;Cross-references: GDB:676446; OMIM:600823
A;Map position: 16p11.2-16p11.2
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; hydrolase; serine proteinase; transmembrane protein
F;1-32/Domain: signal sequence #status predicted <SIG>
F;33-44,45-343/Product: prostasin #status predicted <MAT>
F;33-44/Domain: prostasin light chain #status predicted <CHL>
F;45-343/Domain: prostasin heavy chain #status predicted <CHH>
F;45-281/Domain: trypsin homology <TRY>
F;323-341/Domain: transmembrane #status predicted <TMM1>
F;37-154,70-86,168-244,201-223,234-262/Disulfide bonds: #status predicted
F;85,134,238/Active site: His, Asp, Ser #status predicted
F;159/Binding site: carbohydrate (Asn) (covalent) #status experimental

  Query Match             36.8%;  Score 512;  DB 1;  Length 343;
  Best Local Similarity   41.4%;  Pred. No. 6.7e-40;
  Matches  103;  Conservative   39;  Mismatches   97;  Indels   10;  Gaps    6;

Qy          2 EKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALR 61
              : || ||  |   :||||||:   |   :| |||||:  ||||:|| |  |      |
Db         42 QARITGGSSAVAGQWPWQVSITYEG---VHVCGGSLVSEQWVLSAAHCF-PSEHHKEAYE 97

Qy         62 VQLREQHL--YYQD-QLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPP 118
              |:|    |  | :| ::  :  || ||  :       |||||:|  |:  | ::  : ||
Db         98 VKLGAHQLDSYSEDAKVSTLKDIIPHPSYLQEGSQGDIALLQLSRPITFSRYIRPICLPA 157

Qy        119 ASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRI 178
              |: :|| |: | ||||| |     |  | ||:|::||::    |:  |:: |   ::
Db        158 ANASFPNGLHCTVTGWGHVAPSVSLLTPKPLQQLEVPLISRETCNCLYNIDA-KPEEPHF 216

Qy        179 VRDDMLCAGNTR--RDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTY 236
              |::||:|||     :|:||||||||| | | | |   |:||||: |   ||||:||  :
Db        217 VQEDMVCAGYVEGGKDACQGDSGGPLSCPVEGLWYLTGIVSWGDACGARNRPGVYTLASS 276

Qy        237 YLDWIHHYV 245
              |  ||   |
Db        277 YASWIQSKV 285
```

```
RESULT 14
KFHU1
coagulation factor XIa (EC 3.4.21.27) precursor [validated] - human
N;Alternate names: antihemophilic factor C; plasma thromboplastin antecedent
C;Species: Homo sapiens (man)
C;Date: 13-Aug-1986 #sequence_revision 26-May-1994 #text_change 08-Dec-2000
C;Accession: A27431; A00920; A37940
R;Asakai, R.; Davie, E.W.; Chung, D.W.
Biochemistry 26, 7221-7228, 1987
A;Title: Organization of the gene for human factor XI.
A;Reference number: A27431; MUID:88107663; PMID:2827746
A;Accession: A27431
A;Molecule type: DNA
A;Residues: 1-625 <ASA>
A;Cross-references: GB:M18295
A;Note: the sequence shown follows the authors' translation
R;Fujikawa, K.; Chung, D.W.; Hendrickson, L.E.; Davie, E.W.
Biochemistry 25, 2417-2424, 1986
A;Title: Amino acid sequence of human factor XI, a blood coagulation factor with four tar
A;Reference number: A00920; MUID:86243360; PMID:3636155
A;Accession: A00920
A;Molecule type: mRNA
A;Residues: 1-625 <FUJ>
A;Cross-references: GB:M13142; NID:g182832; PIDN:AAA52487.1; PID:g182833
R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.
Biochemistry 30, 2056-2060, 1991
A;Title: Location of the disulfide bonds in human coagulation factor XI: the presence of
A;Reference number: A37940; MUID:91152017; PMID:1998667
A;Accession: A37940
A;Molecule type: protein
A;Residues: 28-33;35-49,'X',51-55,'X',57-63;70-75,'X',77-79;107-109,'X',111-112;132-139;'
;280-282,'X',284;285-297;313-316,'X',318-319;320-326;'X',328-330;'X',347-349;373,'X',375;
C;Comment: The proenzyme consists of two identical chains linked by one or more disulfid
he active site, and a heavy chain, which associates with high molecular weight (HMW) kini
C;Genetics:
A;Gene: GDB:F11
A;Cross-references: GDB:119891; OMIM:264900
A;Map position: 4q35-4q35
A;Introns: 19/1; 73/2; 109/1; 162/2; 199/1; 252/2; 289/1; 343/2; 379/1; 435/2; 494/1; 52
C;Function:
A;Description: catalyzes the proteolytic activation of coagulation factor IX
A;Pathway: blood coagulation intrinsic pathway
C;Superfamily: coagulation factor XI; trypsin homology
C;Keywords: blood coagulation; duplication; glycoprotein; hemophilia C; homodimer; hydrol
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-387/Product: coagulation factor XIa heavy chain #status experimental <HCH>
F;19-108/Domain: apple repeat <AP1>
F;109-198/Domain: apple repeat <AP2>
F;199-288/Domain: apple repeat <AP3>
F;290-379/Domain: apple repeat <AP4>
F;388-625/Product: coagulation factor XIa light chain #status experimental <LCH>
F;388-618/Domain: trypsin homology <TRY>
F;20-103,514-581,571-599/Disulfide bonds: #status predicted
F;29/Disulfide bonds: interchain #status experimental
F;46-76,50-56,110-193,136-165,140-146,200-283,226-255,230-236,291-374,317-346,321-327,38
F;90,126,353,450/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;339/Disulfide bonds: interchain #status predicted
F;387-388/Cleavage site: Arg-Ile (coagulation factor XIIa) #status experimental
F;431,480,575/Active site: His, Asp, Ser #status predicted
F;491/Binding site: carbohydrate (Asn) (covalent) #status experimental

  Query Match             33.5%;  Score 467;  DB 1;  Length 625;
  Best Local Similarity   39.3%;  Pred. No. 2.2e-35;
  Matches   99;  Conservative   31;  Mismatches   94;  Indels   28;  Gaps    6;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              :: |||||  : | :|||||:|    |   | ||||:|  ||:|||| |    |:    |
Db        384 IKPRIVGGTASVRGEWPWQVTLHTTSPTQRHLCGGSIIGNQWILTAAHCF-YGVESPKIL 442

Qy         61 RVQ---LREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLP 117
              ||    | :  :        |  ||:| |:  |: | |||||:||  |  :     : ||
Db        443 RVYSGILNQSEIKEDTSFFGVQEIIIHDQYKMAESGYDIALLKLETTVNYTDSQRPICLP 502
```

```
Qy          118 PASETFPPGMPCWVTGWG------DVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAY 171
                   :       |||||||       : |         |:: |:|:: |  |  :|
Db          503 SKGDRNVIYTDCWVTGWGYRKLRDKIQN--------TLQKAKIPLVTNEECQKRYRGHKI 554

Qy          172 TGDDVRIVRDDMLCAGNTR--RDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPG 229
                |          |:|||     :|:|:||||||| || |  |   |: ||||||||  |||
Db          555 T--------HKMICAGYREGGKDACKGDSGGPLSCKHNEVWHLVGITSWGEGCAQRERPG 606

Qy          230 IYTRVTYYLDWI 241
                :|| |  |:|||
Db          607 VYTNVVEYVDWI 618

RESULT 15
KQHUP
plasma kallikrein (EC 3.4.21.34) precursor - human
N;Alternate names: kininogenin; plasma prekallikrein
C;Species: Homo sapiens (man)
C;Date: 13-Aug-1986 #sequence_revision 13-Aug-1986 #text_change 18-Jun-1999
C;Accession: A00921; A37939
R;Chung, D.W.; Fujikawa, K.; McMullen, B.A.; Davie, E.W.
Biochemistry 25, 2410-2417, 1986
A;Title: Human plasma prekallikrein, a zymogen to a serine protease that contains four t
A;Reference number: A00921; MUID:86243359; PMID:3521732
A;Accession: A00921
A;Molecule type: mRNA
A;Residues: 1-638 <CHU>
A;Cross-references: GB:M13143; NID:g190262; PIDN:AAA60153.1; PID:g190263
R;McMullen, B.A.; Fujikawa, K.; Davie, E.W.
Biochemistry 30, 2050-2056, 1991
A;Title: Location of the disulfide bonds in human plasma prekallikrein: the presence of
A;Reference number: A37939; MUID:91152016; PMID:1998666
A;Accession: A37939
A;Molecule type: protein
A;Residues: 20-27;40-46,'X',48,'H';50,'X',52-70,'H';75-76,'X',78-80;103-113;131-140;141-
;260-283,'X',285;287-291,'X',293-295;314-317,'X',319-320;321-324;'X',329-333;334-339,'X'
525;538-551;562,'X',564-567;573,'X',575-576;578-583,'X',585;592-604 <MCM>
C;Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w
C;Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a li
 are linked by one or more disulfide bonds.
C;Comment: The enzyme cleaves Lys-Arg and Arg-Ser bonds. It activates, in a reciprocal r
inogen and may also play a role in the renin-angiotensin system by converting prorenin i
C;Genetics:
A;Gene: GDB:KLK3
A;Cross-references: GDB:127575; OMIM:229000
A;Map position: 4q35-4q35
C;Superfamily: coagulation factor XI; trypsin homology
C;Keywords: blood coagulation; duplication; fibrinolysis; glycoprotein; hydrolase; infla
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-638/Product: plasma kallikrein #status predicted <MAT>
F;20-390/Domain: plasma kallikrein heavy chain #status predicted <HCH>
F;20-109/Domain: apple repeat <AP1>
F;110-199/Domain: apple repeat <AP2>
F;200-289/Domain: apple repeat <AP3>
F;291-380/Domain: apple repeat <AP4>
F;391-638/Domain: plasma kallikrein light chain #status predicted <LCH>
F;391-621/Domain: trypsin homology <TRY>
F;21-104,47-77,51-57,111-194,137-166,141-147,201-284,227-256,231-237,292-375,322-328,383
F;127,308,396,453,494/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;318-347,340-345/Disulfide bonds: #status predicted
F;390-391/Cleavage site: Arg-Ile (coagulation factor XIIa) #status predicted
F;434,483,578/Active site: His, Asp, Ser #status predicted

  Query Match             32.4%;  Score 451;  DB 1;  Length 638;
  Best Local Similarity   38.4%;  Pred. No. 6.9e-34;
  Matches   96;  Conservative   40;  Mismatches   84;  Indels   30;  Gaps    8;

Qy            4 RIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACV-GPDVKDLAALRV 62
                |||||  :   :|||||||:|      | ||||||  ||||||| |  |  ::|:   |:
Db          390 RIVGGTNSSWGEWPWQVSLQVKLTAQRHLCGGSLIGHQWVLTAAHCFDGLPLQDV--WRI 447

Qy           63 Q---LREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPA 119
                    |    :        :  ||:|  :  ::   ||||::|: |:  :     : ||
Db          448 YSGILNLSDITKDTPFSQIKEIIIHQNYKVSEGNHDIALIKLQAPLNYTEFQKPICLPSK 507

Qy          120 SETFPPGMPCWVTGW------GDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTG 173
                 :|      ||||||      |:: |         |::| :|:: |  |  :|
Db          508 GDTSTIYTNCWVTGWGFSKEKGEIQN--------ILQKVNIPLVTNEECQKRY------- 552

Qy          174 DDVRIVRDDMLCAGNTR--RDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIY 231
                 | :| :  |:|||     :|:|:|||||||||| || |   |: |||||||:  :||:|
Db          553 QDYKITQ-RMVCAGYKEGGKDACKGDSGGPLVCKHNGMWRLVGITSWGEGCARREQPGVY 611

Qy          232 TRVTYYLDWI 241
                |:|  |:|||
Db          612 TKVAEYMDWI 621

Search completed: July 22, 2004, 15:26:16
Job time : 57 secs
```

09-598-952

```
> O <
O| |O  IntelliGenetics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Fri 23 Jul 104 8:15:25-PDT

Solution Parameters:

    Amino Alphabet      = Identity
    Output line length = 80
    Compress            = Off
    Histogram           = Off
    Randomization       = Off

    AMINO-Res-length    = 2
    DELetion-weight     = 5.00
    LEngth-factor       = 0
    Matching-weight     = 1.00
    NUCLEIC-Res-length = 4
    SPread-factor       = 50

Clustered order of selected sequences:

    16. US-09-598-982-43     (1-249)
    15. US-09-598-982-41     (1-249)
     7. US-09-598-982-25     (1-249)
     8. US-09-598-982-27     (1-249)
     3. US-09-598-982-9      (1-249)
     5. US-09-598-982-21     (1-249)
     6. US-09-598-982-23     (1-249)
    14. US-09-598-982-39     (1-249)
    13. US-09-598-982-37     (1-249)
     4. US-09-598-982-11     (1-245)

Region Alignment: (listed in Clustered order)

US-09-598-     1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALR
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-     1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALR
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-     1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALR
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-     1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALR
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-     1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALR
                 ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
US-09-598-     1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAaCVGPDVKDLAALR
                 ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
US-09-598-     1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALR
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-     1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALR
                 ||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
US-09-598-     1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAaCVGPDVKDLAALR
                     ||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
US-09-598-     1     IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAhCVGPDVKDLAALR

 consensus       lekrIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAhCVGPDVKDLAALR

US-09-598-    62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASET
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASET
                 |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
US-09-598-    62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASET
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASET
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASET
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASET
                 ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
US-09-598-    62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAAIALLELEEPVKVSSHVHTVTLPPASET
                 |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
US-09-598-    62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAAIALLELEEPVNVSSHVHTVTLPPASET
                 ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
US-09-598-    62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASET
                 |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
US-09-598-    58 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVkVSSHVHTVTLPPASET

 consensus       VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAdIALLELEEPVkVSSHVHTVTLPPASET

US-09-598-   123 FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   123 FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   123 FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   123 FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   123 FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   123 FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   123 FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   123 FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   123 FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   119 FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM

 consensus       FPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDM

US-09-598-   184 LCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   184 LCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   184 LCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   184 LCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY
                 |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   184 LCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   184 LCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   184 LCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   184 LCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   184 LCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   180 LCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY

 consensus       LCAGNTRRDSCQGDsGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY

US-09-598-   245 VPKKP
                 |||||
US-09-598-   245 VPKKP
                 |||||
US-09-598-   245 VPKKP
                 |||||
US-09-598-   245 VPKKP
                 |||||
US-09-598-   245 VPKKP
                 |||||
US-09-598-   245 VPKKP
                 |||||
US-09-598-   245 VPKKP
                 |||||
```

```
US-09-598-  245 VPKKP
                |||||
US-09-598-  245 VPKKP
                |||||
US-09-598-  241 VPKKP

 consensus      VPKKP

Alignment score = 2213.00

Scoring matrix:

         3    4    5    6    7    8   13   14   15   16
      ---------------------------------------------------
  3  |        225  248  248  248  248  247  247  247  247
  4  |             224  224  224  224  223  223  223  223
  5  |                  247  247  247  248  246  246  246
  6  |                       247  247  246  248  246  246
  7  |                            249  246  246  248  248
  8  |                                 246  246  248  248
 13  |                                      247  247  247
 14  |                                           247  247
 15  |                                                249
 16  |

-------
```

```
> O <
O| |O  IntelliGenetics
> O <

GENALIGN - Multiple Sequence Alignment Program
Release 5.4

Fri 23 Jul 104 8:12:22-PDT

Solution Parameters:

    Nucleic Alphabet    = Identity
    Output line length = 80
    Compress            = Off
    Histogram           = Off
    Randomization       = Off

    AMINO-Res-length    = 2
    DELetion-weight     = 5.00
    LEngth-factor       = 0
    Matching-weight     = 1.00
    NUCLEIC-Res-length = 4
    SPread-factor       = 50

Clustered order of selected sequences:

    27. US-09-598-982-42      (1-771)
    19. US-09-598-982-26      (1-771)
    18. US-09-598-982-24      (1-771)
    26. US-09-598-982-40      (1-771)
    25. US-09-598-982-38      (1-771)
    17. US-09-598-982-22      (1-771)
     6. US-09-598-982-8       (1-771)
    16. US-09-598-982-20      (1-771)
    24. US-09-598-982-36      (1-771)
     7. US-09-598-982-10      (1-735)

Region Alignment: (listed in Clustered order)

US-09-598-    1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-    1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC
                                  |||||||||||||||||||||||||||||||||||||||||||
US-09-598-    1                   ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC

 consensus      gggcccctcgagaaaagaATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGC


US-09-598-   62 AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   62 AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   62 AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   62 AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   62 AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   62 AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   62 AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-✓  62 AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   62 AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   44 AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA

 consensus      AGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCA


US-09-598-  123 CCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  123 CCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  123 CCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  123 CCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  123 CCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  123 CCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  123 CCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC
                ||||||||||||||||||||| |||   |||||||||||||||||||||||||||||||||
US-09-598-✓ 123 CCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  123 CCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC
                ||||||||||||||||||||| |||   |||||||||||||||||||||||||||||||||
US-09-598-  105 CCCCCAGTGGGTGCTGACCGCaGCGcacTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC

 consensus      CCCCCAGTGGGTGCTGACCGCaGCGcacTGCGTGGGACCGGACGTCAAGGATCTGGCCGCC


US-09-598-  184 CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  184 CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  184 CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  184 CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  184 CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  184 CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  184 CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-✓ 184 CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  184 CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  166 CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA

 consensus      CTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCA


US-09-598-  245 GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  245 GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  245 GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  245 GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGA
                |||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
US-09-598-  245 GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTGCTGGA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  245 GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTGCTGGA
                |||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||
US-09-598-  245 GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
US-09-598-√ 245 GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  245 GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  227 GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGA

 consensus      GGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGacATCGCCCTGCTGGA

US-09-598-  306 GCTGGAGGAGCCGGTGAAcGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA
                |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
US-09-598-  306 GCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  306 GCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA
                |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
US-09-598-  306 GCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  306 GCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA
                |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
US-09-598-  306 GCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  306 GCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-√ 306 GCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA
                |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
US-09-598-  306 GCTGGAGGAGCCGGTGAAcGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA
                |||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||
US-09-598-  288 GCTGGAGGAGCCGGTGAAgGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA

 consensus      GCTGGAGGAGCCGGTGAAgGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA

US-09-598-  367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-√ 367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  349 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG

 consensus      GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG

US-09-598-  428 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  428 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  428 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  428 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  428 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  428 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  428 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-√ 428 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  428 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  410 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT

 consensus      AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACAT

US-09-598-  489 TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  489 TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  489 TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  489 TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  489 TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  489 TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  489 TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-√ 489 TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  489 TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  471 TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC

 consensus      TTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGAC

US-09-598-  550 GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCCGGCGGACCAC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  550 GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCCGGCGGACCAC
                |||||||||||||||||||||||||||||||||||||| ||||| |||||||||||||| |
US-09-598-  550 GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCCGGCGGACCTC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  550 GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCCGGCGGACCTC
                |||||||||||||||||||||||||||||||||||||| || |||||| |||| || || |
US-09-598-  550 GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  550 GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  550 GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-√ 550 GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  550 GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCC
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  532 GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCC

 consensus      GACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGcCAgGGcGACtCCGGaGGgCCcC

US-09-598-  611 TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  611 TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  611 TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  611 TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  611 TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  611 TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  611 TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-√ 611 TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  611 TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-  593 TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG

 consensus      TGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTG
```

```
US-09-598-   672 TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   672 TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   672 TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   672 TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   672 TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   672 TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   672 TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-✓  672 TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   672 TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
US-09-598-   654 TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC

 consensus       TGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCAC

US-09-598-   733 CACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT
                 |||||||||||||||||||||||||||||||||||||||
US-09-598-   733 CACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT
                 |||||||||||||||||||||||||||||||||||||||
US-09-598-   733 CACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT
                 |||||||||||||||||||||||||||||||||||||||
US-09-598-   733 CACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT
                 |||||||||||||||||||||||||||||||||||||||
US-09-598-   733 CACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT
                 |||||||||||||||||||||||||||||||||||||||
US-09-598-   733 CACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT
                 |||||||||||||||||||||||||||||||||||||||
US-09-598-   733 CACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT
                 |||||||||||||||||||||||||||||||||||||||
US-09-598-✓  733 CACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT
                 |||||||||||||||||||||||||||||||||||||||
US-09-598-   733 CACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT
                 |||||||||||||||||||||
US-09-598-   715 CACTATGTCCCCAAAAAGCCG

 consensus       CACTATGTCCCCAAAAAGCCGtgaagcggccgccgtcgt

Alignment score = 6708.00

Scoring matrix:

        6     7    16    17    18    19    24    25    26    27
    ------------------------------------------------------------
  6  |       555   767   769   765   765   766   768   764   764
  7  |             551   553   549   549   550   552   548   548
 16  |                   765   761   761   770   764   760   760
 17  |                         763   763   764   770   762   762
 18  |                               768   760   762   770   767
 19  |                                     760   762   767   770
 24  |                                           765   761   761
 25  |                                                 763   763
 26  |                                                       768

       27 |

    -------
```

```
                     GenCore version 5.1.6
            Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 23, 2004, 08:23:52 ; Search time 0.001 Seconds
                                          (without alignments)
                                          383.958 Million cell updates/sec

Title:          US-09-598-982-9
Perfect score:  1397
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1 seqs, 771 residues

Total number of hits satisfying chosen parameters:      2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=Pending_Patents_AA_Main:US-09-598-982-9
-DB=Pending_Patents_NA_Main:US-09-598-982-8 -SUFFIX=pto -OUT=align9_8
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Pending_Patents_NA_Main:US-09-598-982-8
                1:  /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:US-09-598-982-8
                2:  /cgn2_6/ptodata/2/pna/US06_COMB.seq:US-09-598-982-8
                3:  /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-09-598-982-8
                4:  /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-09-598-982-8
                5:  /cgn2_6/ptodata/2/pna/US081_COMB.seq:US-09-598-982-8
                6:  /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-09-598-982-8
                7:  /cgn2_6/ptodata/2/pna/US083_COMB.seq:US-09-598-982-8
                8:  /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-09-598-982-8
                9:  /cgn2_6/ptodata/2/pna/US085_COMB.seq:US-09-598-982-8
                10:  /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-09-598-982-8
                11:  /cgn2_6/ptodata/2/pna/US087_COMB.seq:US-09-598-982-8
                12:  /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-09-598-982-8
                13:  /cgn2_6/ptodata/2/pna/US089_COMB.seq:US-09-598-982-8
                14:  /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-598-982-8
                15:  /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-598-982-8
                16:  /cgn2_6/ptodata/2/pna/US092A_COMB.seq:US-09-598-982-8
                17:  /cgn2_6/ptodata/2/pna/US092B_COMB.seq:US-09-598-982-8
                18:  /cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-598-982-8
                19:  /cgn2_6/ptodata/2/pna/US093B_COMB.seq:US-09-598-982-8
                20:  /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-598-982-8
                21:  /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-598-982-8
                22:  /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-598-982-8
                23:  /cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-598-982-8
                24:  /cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-598-982-8
                25:  /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-598-982-8
                26:  /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-598-982-8
                27:  /cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-598-982-8
                28:  /cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-598-982-8
                29:  /cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-598-982-8
                30:  /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-598-982-8
                31:  /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-598-982-8
                32:  /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-598-982-8
                33:  /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-598-982-8
                34:  /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-598-982-8
                35:  /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-598-982-8
                36:  /cgn2_6/ptodata/2/pna/US098D_COMB.seq:US-09-598-982-8
                37:  /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-598-982-8
                38:  /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-598-982-8
                39:  /cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-598-982-8
                40:  /cgn2_6/ptodata/2/pna/US099D_COMB.seq:US-09-598-982-8
                41:  /cgn2_6/ptodata/2/pna/US099E_COMB.seq:US-09-598-982-8
                42:  /cgn2_6/ptodata/2/pna/US099F_COMB.seq:US-09-598-982-8
                43:  /cgn2_6/ptodata/2/pna/US100A_COMB.seq:US-09-598-982-8
                44:  /cgn2_6/ptodata/2/pna/US100B_COMB.seq:US-09-598-982-8
                45:  /cgn2_6/ptodata/2/pna/US101A_COMB.seq:US-09-598-982-8
                46:  /cgn2_6/ptodata/2/pna/US101B_COMB.seq:US-09-598-982-8
                47:  /cgn2_6/ptodata/2/pna/US102A_COMB.seq:US-09-598-982-8
                48:  /cgn2_6/ptodata/2/pna/US102B_COMB.seq:US-09-598-982-8
                49:  /cgn2_6/ptodata/2/pna/US103A_COMB.seq:US-09-598-982-8
                50:  /cgn2_6/ptodata/2/pna/US103B_COMB.seq:US-09-598-982-8
                51:  /cgn2_6/ptodata/2/pna/US104A_COMB.seq:US-09-598-982-8
                52:  /cgn2_6/ptodata/2/pna/US104B_COMB.seq:US-09-598-982-8
                53:  /cgn2_6/ptodata/2/pna/US106_COMB.seq:US-09-598-982-8
                54:  /cgn2_6/ptodata/2/pna/US107A_COMB.seq:US-09-598-982-8
                55:  /cgn2_6/ptodata/2/pna/US107B_COMB.seq:US-09-598-982-8
                56:  /cgn2_6/ptodata/2/pna/US6000_COMB.seq:US-09-598-982-8
                57:  /cgn2_6/ptodata/2/pna/US6001_COMB.seq:US-09-598-982-8
                58:  /cgn2_6/ptodata/2/pna/US6002_COMB.seq:US-09-598-982-8
                59:  /cgn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-598-982-8
                60:  /cgn2_6/ptodata/2/pna/US6004_COMB.seq:US-09-598-982-8
                61:  /cgn2_6/ptodata/2/pna/US6005_COMB.seq:US-09-598-982-8
                62:  /cgn2_6/ptodata/2/pna/US6006_COMB.seq:US-09-598-982-8
                63:  /cgn2_6/ptodata/2/pna/US6007_COMB.seq:US-09-598-982-8
                64:  /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-09-598-982-8
                65:  /cgn2_6/ptodata/2/pna/US6009_COMB.seq:US-09-598-982-8
                66:  /cgn2_6/ptodata/2/pna/US6010_COMB.seq:US-09-598-982-8
                67:  /cgn2_6/ptodata/2/pna/US6011_COMB.seq:US-09-598-982-8
                68:  /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-09-598-982-8
                69:  /cgn2_6/ptodata/2/pna/US6013_COMB.seq:US-09-598-982-8
                70:  /cgn2_6/ptodata/2/pna/US6014_COMB.seq:US-09-598-982-8
                71:  /cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-09-598-982-8
                72:  /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-09-598-982-8
                73:  /cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-09-598-982-8
                74:  /cgn2_6/ptodata/2/pna/US6018_COMB.seq:US-09-598-982-8
                75:  /cgn2_6/ptodata/2/pna/US6019_COMB.seq:US-09-598-982-8
                76:  /cgn2_6/ptodata/2/pna/US6020_COMB.seq:US-09-598-982-8
                77:  /cgn2_6/ptodata/2/pna/US6021_COMB.seq:US-09-598-982-8
                78:  /cgn2_6/ptodata/2/pna/US6022_COMB.seq:US-09-598-982-8
                79:  /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:US-09-598-982-8
                80:  /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:US-09-598-982-8
                81:  /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-09-598-982-8
                82:  /cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-09-598-982-8
                83:  /cgn2_6/ptodata/2/pna/US6026_COMB.seq:US-09-598-982-8
                84:  /cgn2_6/ptodata/2/pna/US6027_COMB.seq:US-09-598-982-8
                85:  /cgn2_6/ptodata/2/pna/US6028_COMB.seq:US-09-598-982-8
                86:  /cgn2_6/ptodata/2/pna/US6029_COMB.seq:US-09-598-982-8
                87:  /cgn2_6/ptodata/2/pna/US6030_COMB.seq:US-09-598-982-8
                88:  /cgn2_6/ptodata/2/pna/US6031_COMB.seq:US-09-598-982-8
                89:  /cgn2_6/ptodata/2/pna/US6032_COMB.seq:US-09-598-982-8
                90:  /cgn2_6/ptodata/2/pna/US6033_COMB.seq:US-09-598-982-8
                91:  /cgn2_6/ptodata/2/pna/US6034_COMB.seq:US-09-598-982-8
                92:  /cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-09-598-982-8
                93:  /cgn2_6/ptodata/2/pna/US6036_COMB.seq:US-09-598-982-8
                94:  /cgn2_6/ptodata/2/pna/US6037_COMB.seq:US-09-598-982-8
                95:  /cgn2_6/ptodata/2/pna/US6038_COMB.seq:US-09-598-982-8
                96:  /cgn2_6/ptodata/2/pna/US6039_COMB.seq:US-09-598-982-8
                97:  /cgn2_6/ptodata/2/pna/US6040_COMB.seq:US-09-598-982-8
                98:  /cgn2_6/ptodata/2/pna/US6041_COMB.seq:US-09-598-982-8
                99:  /cgn2_6/ptodata/2/pna/US6042_COMB.seq:US-09-598-982-8
                100:  /cgn2_6/ptodata/2/pna/US6043_COMB.seq:US-09-598-982-8
                101:  /cgn2_6/ptodata/2/pna/US6044_COMB.seq:US-09-598-982-8
                102:  /cgn2_6/ptodata/2/pna/US6045_COMB.seq:US-09-598-982-8
                103:  /cgn2_6/ptodata/2/pna/US6046_COMB.seq:US-09-598-982-8
```

```
           104:  /cgn2_6/ptodata/2/pna/US6047_COMB.seq:US-09-598-982-8
           105:  /cgn2_6/ptodata/2/pna/US6048_COMB.seq:US-09-598-982-8
           106:  /cgn2_6/ptodata/2/pna/US6049_COMB.seq:US-09-598-982-8
           107:  /cgn2_6/ptodata/2/pna/US6050_COMB.seq:US-09-598-982-8
           108:  /cgn2_6/ptodata/2/pna/US6051_COMB.seq:US-09-598-982-8
           109:  /cgn2_6/ptodata/2/pna/US6052_COMB.seq:US-09-598-982-8
           110:  /cgn2_6/ptodata/2/pna/US6053_COMB.seq:US-09-598-982-8

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                   %
Result           Query
  No.    Score   Match Length DB  ID                        Description
-------------------------------------------------------------------------------
    1     1397   100.0    771  24  US-09-598-982-8           Sequence 8, Appli

                                   ALIGNMENTS

RESULT 1
US-09-598-982-8
; Sequence 8, Application US/09598982
; GENERAL INFORMATION:
;  APPLICANT:  NILES, ANDREW L
;  APPLICANT:  MAFFITT, MARK A
;  APPLICANT:  HAAK-FRENDSCHO, MARY
;  TITLE OF INVENTION:  RECOMBINANT PROTEOLYTIC TYPTASES, ACTIVE SITE MUTANTS
;  TITLE OF INVENTION:  THEREOF, AND METHODS OF MAKING SAME
;  FILE REFERENCE:  CIP TRYPTASE
;  CURRENT APPLICATION NUMBER: US/09/598,982
;  CURRENT FILING DATE:  2000-06-21
;  PRIOR APPLICATION NUMBER:  09/079,970
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  51
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 8
;    LENGTH: 771
;    TYPE: DNA
;    ORGANISM: Homo sapiens
;    FEATURE:
;    NAME/KEY: CDS
;    LOCATION: (7)..(753)
US-09-598-982-8

 Alignment Scores:
 Pred. No.:               0              Length:          771
 Score:                   1397.00        Matches:         249
 Percent Similarity:      100.00%        Conservative:    0
 Best Local Similarity:   100.00%        Mismatches:      0
 Query Match:             100.00%        Indels:          0
 DB:                      24             Gaps:            0

US-09-598-982-9 (1-249) x US-09-598-982-8 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246
```

```
Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

Search completed: July 23, 2004, 08:23:53
Job time : 0.001 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 23, 2004, 08:25:03 ; Search time 1 Seconds
                                           (without alignments)
                                           0.360 Million cell updates/sec

Title:          US-09-598-982-11
Perfect score:  1378
Sequence:       1 IVGGQEAPRSKWPWQVSLRV..........IYTRVTYYLDWIHHYVPKKP 245

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1 seqs, 735 residues

Total number of hits satisfying chosen parameters:      2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=Pending_Patents_AA_Main:US-09-598-982-11
-DB=Pending_Patents_NA_Main:US-09-598-982-10 -SUFFIX=pto -OUT=align11_10
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Pending_Patents_NA_Main:US-09-598-982-10
                1:  /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:US-09-598-982-10
                2:  /cgn2_6/ptodata/2/pna/US06_COMB.seq:US-09-598-982-10
                3:  /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-09-598-982-10
                4:  /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-09-598-982-10
                5:  /cgn2_6/ptodata/2/pna/US081_COMB.seq:US-09-598-982-10
                6:  /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-09-598-982-10
                7:  /cgn2_6/ptodata/2/pna/US083_COMB.seq:US-09-598-982-10
                8:  /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-09-598-982-10
                9:  /cgn2_6/ptodata/2/pna/US085_COMB.seq:US-09-598-982-10
                10:  /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-09-598-982-10
                11:  /cgn2_6/ptodata/2/pna/US087_COMB.seq:US-09-598-982-10
                12:  /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-09-598-982-10
                13:  /cgn2_6/ptodata/2/pna/US089_COMB.seq:US-09-598-982-10
                14:  /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-598-982-10
                15:  /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-598-982-10
                16:  /cgn2_6/ptodata/2/pna/US092A_COMB.seq:US-09-598-982-10
                17:  /cgn2_6/ptodata/2/pna/US092B_COMB.seq:US-09-598-982-10
                18:  /cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-598-982-10
                19:  /cgn2_6/ptodata/2/pna/US093B_COMB.seq:US-09-598-982-10
                20:  /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-598-982-10
                21:  /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-598-982-10
                22:  /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-598-982-10
                23:  /cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-598-982-10
                24:  /cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-598-982-10
                25:  /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-598-982-10
                26:  /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-598-982-10
                27:  /cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-598-982-10
                28:  /cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-598-982-10
                29:  /cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-598-982-10
                30:  /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-598-982-10
```

```
31:  /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-598-982-10
32:  /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-598-982-10
33:  /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-598-982-10
34:  /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-598-982-10
35:  /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-598-982-10
36:  /cgn2_6/ptodata/2/pna/US098D_COMB.seq:US-09-598-982-10
37:  /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-598-982-10
38:  /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-598-982-10
39:  /cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-598-982-10
40:  /cgn2_6/ptodata/2/pna/US099D_COMB.seq:US-09-598-982-10
41:  /cgn2_6/ptodata/2/pna/US099E_COMB.seq:US-09-598-982-10
42:  /cgn2_6/ptodata/2/pna/US099F_COMB.seq:US-09-598-982-10
43:  /cgn2_6/ptodata/2/pna/US100A_COMB.seq:US-09-598-982-10
44:  /cgn2_6/ptodata/2/pna/US100B_COMB.seq:US-09-598-982-10
45:  /cgn2_6/ptodata/2/pna/US101A_COMB.seq:US-09-598-982-10
46:  /cgn2_6/ptodata/2/pna/US101B_COMB.seq:US-09-598-982-10
47:  /cgn2_6/ptodata/2/pna/US102A_COMB.seq:US-09-598-982-10
48:  /cgn2_6/ptodata/2/pna/US102B_COMB.seq:US-09-598-982-10
49:  /cgn2_6/ptodata/2/pna/US103A_COMB.seq:US-09-598-982-10
50:  /cgn2_6/ptodata/2/pna/US103B_COMB.seq:US-09-598-982-10
51:  /cgn2_6/ptodata/2/pna/US104A_COMB.seq:US-09-598-982-10
52:  /cgn2_6/ptodata/2/pna/US104B_COMB.seq:US-09-598-982-10
53:  /cgn2_6/ptodata/2/pna/US106_COMB.seq:US-09-598-982-10
54:  /cgn2_6/ptodata/2/pna/US107A_COMB.seq:US-09-598-982-10
55:  /cgn2_6/ptodata/2/pna/US107B_COMB.seq:US-09-598-982-10
56:  /cgn2_6/ptodata/2/pna/US6000_COMB.seq:US-09-598-982-10
57:  /cgn2_6/ptodata/2/pna/US6001_COMB.seq:US-09-598-982-10
58:  /cgn2_6/ptodata/2/pna/US6002_COMB.seq:US-09-598-982-10
59:  /cgn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-598-982-10
60:  /cgn2_6/ptodata/2/pna/US6004_COMB.seq:US-09-598-982-10
61:  /cgn2_6/ptodata/2/pna/US6005_COMB.seq:US-09-598-982-10
62:  /cgn2_6/ptodata/2/pna/US6006_COMB.seq:US-09-598-982-10
63:  /cgn2_6/ptodata/2/pna/US6007_COMB.seq:US-09-598-982-10
64:  /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-09-598-982-10
65:  /cgn2_6/ptodata/2/pna/US6009_COMB.seq:US-09-598-982-10
66:  /cgn2_6/ptodata/2/pna/US6010_COMB.seq:US-09-598-982-10
67:  /cgn2_6/ptodata/2/pna/US6011_COMB.seq:US-09-598-982-10
68:  /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-09-598-982-10
69:  /cgn2_6/ptodata/2/pna/US6013_COMB.seq:US-09-598-982-10
70:  /cgn2_6/ptodata/2/pna/US6014_COMB.seq:US-09-598-982-10
71:  /cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-09-598-982-10
72:  /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-09-598-982-10
73:  /cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-09-598-982-10
74:  /cgn2_6/ptodata/2/pna/US6018_COMB.seq:US-09-598-982-10
75:  /cgn2_6/ptodata/2/pna/US6019_COMB.seq:US-09-598-982-10
76:  /cgn2_6/ptodata/2/pna/US6020_COMB.seq:US-09-598-982-10
77:  /cgn2_6/ptodata/2/pna/US6021_COMB.seq:US-09-598-982-10
78:  /cgn2_6/ptodata/2/pna/US6022_COMB.seq:US-09-598-982-10
79:  /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:US-09-598-982-10
80:  /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:US-09-598-982-10
81:  /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-09-598-982-10
82:  /cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-09-598-982-10
83:  /cgn2_6/ptodata/2/pna/US6026_COMB.seq:US-09-598-982-10
84:  /cgn2_6/ptodata/2/pna/US6027_COMB.seq:US-09-598-982-10
85:  /cgn2_6/ptodata/2/pna/US6028_COMB.seq:US-09-598-982-10
86:  /cgn2_6/ptodata/2/pna/US6029_COMB.seq:US-09-598-982-10
87:  /cgn2_6/ptodata/2/pna/US6030_COMB.seq:US-09-598-982-10
88:  /cgn2_6/ptodata/2/pna/US6031_COMB.seq:US-09-598-982-10
89:  /cgn2_6/ptodata/2/pna/US6032_COMB.seq:US-09-598-982-10
90:  /cgn2_6/ptodata/2/pna/US6033_COMB.seq:US-09-598-982-10
91:  /cgn2_6/ptodata/2/pna/US6034_COMB.seq:US-09-598-982-10
92:  /cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-09-598-982-10
93:  /cgn2_6/ptodata/2/pna/US6036_COMB.seq:US-09-598-982-10
94:  /cgn2_6/ptodata/2/pna/US6037_COMB.seq:US-09-598-982-10
95:  /cgn2_6/ptodata/2/pna/US6038_COMB.seq:US-09-598-982-10
96:  /cgn2_6/ptodata/2/pna/US6039_COMB.seq:US-09-598-982-10
97:  /cgn2_6/ptodata/2/pna/US6040_COMB.seq:US-09-598-982-10
98:  /cgn2_6/ptodata/2/pna/US6041_COMB.seq:US-09-598-982-10
99:  /cgn2_6/ptodata/2/pna/US6042_COMB.seq:US-09-598-982-10
100:  /cgn2_6/ptodata/2/pna/US6043_COMB.seq:US-09-598-982-10
101:  /cgn2_6/ptodata/2/pna/US6044_COMB.seq:US-09-598-982-10
102:  /cgn2_6/ptodata/2/pna/US6045_COMB.seq:US-09-598-982-10
103:  /cgn2_6/ptodata/2/pna/US6046_COMB.seq:US-09-598-982-10
```

```
              104:   /cgn2_6/ptodata/2/pna/US6047_COMB.seq:US-09-598-982-10
              105:   /cgn2_6/ptodata/2/pna/US6048_COMB.seq:US-09-598-982-10
              106:   /cgn2_6/ptodata/2/pna/US6049_COMB.seq:US-09-598-982-10
              107:   /cgn2_6/ptodata/2/pna/US6050_COMB.seq:US-09-598-982-10
              108:   /cgn2_6/ptodata/2/pna/US6051_COMB.seq:US-09-598-982-10
              109:   /cgn2_6/ptodata/2/pna/US6052_COMB.seq:US-09-598-982-10
              110:   /cgn2_6/ptodata/2/pna/US6053_COMB.seq:US-09-598-982-10

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                        SUMMARIES
                       %
Result                Query
  No.    Score    Match Length DB  ID                        Description
-----------------------------------------------------------------------------
    1     1378   100.0    735  24  US-09-598-982-10           Sequence 10, Appl


                                        ALIGNMENTS

RESULT 1
US-09-598-982-10
; Sequence 10, Application US/09598982
; GENERAL INFORMATION:
;  APPLICANT:  NILES, ANDREW L
;  APPLICANT:  MAFFITT, MARK A
;  APPLICANT:  HAAK-FRENDSCHO, MARY
;  TITLE OF INVENTION:  RECOMBINANT PROTEOLYTIC TYPTASES, ACTIVE SITE MUTANTS
;  TITLE OF INVENTION:  THEREOF, AND METHODS OF MAKING SAME
;  FILE REFERENCE:  CIP TRYPTASE
;  CURRENT APPLICATION NUMBER: US/09/598,982
;  CURRENT FILING DATE:  2000-06-21
;  PRIOR APPLICATION NUMBER:  09/079,970
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  51
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 10
;   LENGTH: 735
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(735)
US-09-598-982-10

Alignment Scores:
Pred. No.:               0              Length:        735
Score:                   1378.00        Matches:       245
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             100.00%        Indels:        0
DB:                      24             Gaps:          0

US-09-598-982-11 (1-245) x US-09-598-982-10 (1-735)

Qy          1 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy         21 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy         41 ThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy         61 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240
```

```
Qy         81 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy        101 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy        121 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy        141 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy        161 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy        181 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy        201 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy        221 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy        241 ValProLysLysPro 245
              |||||||||||||||
Db        721 GTCCCCAAAAAGCCG 735


Search completed: July 23, 2004, 08:25:04
Job time : 1 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 23, 2004, 08:27:03 ; Search time 1 Seconds
                                           (without alignments)
                                           0.384 Million cell updates/sec

Title:          US-09-598-982-21
Perfect score:  1393
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1 seqs, 771 residues

Total number of hits satisfying chosen parameters:      2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=Pending_Patents_AA_Main:US-09-598-982-21
-DB=Pending_Patents_NA_Main:US-09-598-982-20 -SUFFIX=pto -OUT=align21_20
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Pending_Patents_NA_Main:US-09-598-982-20
                1:  /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:US-09-598-982-20
                2:  /cgn2_6/ptodata/2/pna/US06_COMB.seq:US-09-598-982-20
                3:  /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-09-598-982-20
                4:  /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-09-598-982-20
                5:  /cgn2_6/ptodata/2/pna/US081_COMB.seq:US-09-598-982-20
                6:  /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-09-598-982-20
                7:  /cgn2_6/ptodata/2/pna/US083_COMB.seq:US-09-598-982-20
                8:  /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-09-598-982-20
                9:  /cgn2_6/ptodata/2/pna/US085_COMB.seq:US-09-598-982-20
                10:  /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-09-598-982-20
                11:  /cgn2_6/ptodata/2/pna/US087_COMB.seq:US-09-598-982-20
                12:  /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-09-598-982-20
                13:  /cgn2_6/ptodata/2/pna/US089_COMB.seq:US-09-598-982-20
                14:  /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-598-982-20
                15:  /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-598-982-20
                16:  /cgn2_6/ptodata/2/pna/US092A_COMB.seq:US-09-598-982-20
                17:  /cgn2_6/ptodata/2/pna/US092B_COMB.seq:US-09-598-982-20
                18:  /cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-598-982-20
                19:  /cgn2_6/ptodata/2/pna/US093B_COMB.seq:US-09-598-982-20
                20:  /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-598-982-20
                21:  /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-598-982-20
                22:  /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-598-982-20
                23:  /cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-598-982-20
                24:  /cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-598-982-20
                25:  /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-598-982-20
                26:  /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-598-982-20
                27:  /cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-598-982-20
                28:  /cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-598-982-20
                29:  /cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-598-982-20
                30:  /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-598-982-20
                31:  /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-598-982-20
                32:  /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-598-982-20
                33:  /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-598-982-20
                34:  /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-598-982-20
                35:  /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-598-982-20
                36:  /cgn2_6/ptodata/2/pna/US098D_COMB.seq:US-09-598-982-20
                37:  /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-598-982-20
                38:  /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-598-982-20
                39:  /cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-598-982-20
                40:  /cgn2_6/ptodata/2/pna/US099D_COMB.seq:US-09-598-982-20
                41:  /cgn2_6/ptodata/2/pna/US099E_COMB.seq:US-09-598-982-20
                42:  /cgn2_6/ptodata/2/pna/US099F_COMB.seq:US-09-598-982-20
                43:  /cgn2_6/ptodata/2/pna/US100A_COMB.seq:US-09-598-982-20
                44:  /cgn2_6/ptodata/2/pna/US100B_COMB.seq:US-09-598-982-20
                45:  /cgn2_6/ptodata/2/pna/US101A_COMB.seq:US-09-598-982-20
                46:  /cgn2_6/ptodata/2/pna/US101B_COMB.seq:US-09-598-982-20
                47:  /cgn2_6/ptodata/2/pna/US102A_COMB.seq:US-09-598-982-20
                48:  /cgn2_6/ptodata/2/pna/US102B_COMB.seq:US-09-598-982-20
                49:  /cgn2_6/ptodata/2/pna/US103A_COMB.seq:US-09-598-982-20
                50:  /cgn2_6/ptodata/2/pna/US103B_COMB.seq:US-09-598-982-20
                51:  /cgn2_6/ptodata/2/pna/US104A_COMB.seq:US-09-598-982-20
                52:  /cgn2_6/ptodata/2/pna/US104B_COMB.seq:US-09-598-982-20
                53:  /cgn2_6/ptodata/2/pna/US106_COMB.seq:US-09-598-982-20
                54:  /cgn2_6/ptodata/2/pna/US107A_COMB.seq:US-09-598-982-20
                55:  /cgn2_6/ptodata/2/pna/US107B_COMB.seq:US-09-598-982-20
                56:  /cgn2_6/ptodata/2/pna/US6000_COMB.seq:US-09-598-982-20
                57:  /cgn2_6/ptodata/2/pna/US6001_COMB.seq:US-09-598-982-20
                58:  /cgn2_6/ptodata/2/pna/US6002_COMB.seq:US-09-598-982-20
                59:  /cgn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-598-982-20
                60:  /cgn2_6/ptodata/2/pna/US6004_COMB.seq:US-09-598-982-20
                61:  /cgn2_6/ptodata/2/pna/US6005_COMB.seq:US-09-598-982-20
                62:  /cgn2_6/ptodata/2/pna/US6006_COMB.seq:US-09-598-982-20
                63:  /cgn2_6/ptodata/2/pna/US6007_COMB.seq:US-09-598-982-20
                64:  /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-09-598-982-20
                65:  /cgn2_6/ptodata/2/pna/US6009_COMB.seq:US-09-598-982-20
                66:  /cgn2_6/ptodata/2/pna/US6010_COMB.seq:US-09-598-982-20
                67:  /cgn2_6/ptodata/2/pna/US6011_COMB.seq:US-09-598-982-20
                68:  /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-09-598-982-20
                69:  /cgn2_6/ptodata/2/pna/US6013_COMB.seq:US-09-598-982-20
                70:  /cgn2_6/ptodata/2/pna/US6014_COMB.seq:US-09-598-982-20
                71:  /cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-09-598-982-20
                72:  /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-09-598-982-20
                73:  /cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-09-598-982-20
                74:  /cgn2_6/ptodata/2/pna/US6018_COMB.seq:US-09-598-982-20
                75:  /cgn2_6/ptodata/2/pna/US6019_COMB.seq:US-09-598-982-20
                76:  /cgn2_6/ptodata/2/pna/US6020_COMB.seq:US-09-598-982-20
                77:  /cgn2_6/ptodata/2/pna/US6021_COMB.seq:US-09-598-982-20
                78:  /cgn2_6/ptodata/2/pna/US6022_COMB.seq:US-09-598-982-20
                79:  /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:US-09-598-982-20
                80:  /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:US-09-598-982-20
                81:  /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-09-598-982-20
                82:  /cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-09-598-982-20
                83:  /cgn2_6/ptodata/2/pna/US6026_COMB.seq:US-09-598-982-20
                84:  /cgn2_6/ptodata/2/pna/US6027_COMB.seq:US-09-598-982-20
                85:  /cgn2_6/ptodata/2/pna/US6028_COMB.seq:US-09-598-982-20
                86:  /cgn2_6/ptodata/2/pna/US6029_COMB.seq:US-09-598-982-20
                87:  /cgn2_6/ptodata/2/pna/US6030_COMB.seq:US-09-598-982-20
                88:  /cgn2_6/ptodata/2/pna/US6031_COMB.seq:US-09-598-982-20
                89:  /cgn2_6/ptodata/2/pna/US6032_COMB.seq:US-09-598-982-20
                90:  /cgn2_6/ptodata/2/pna/US6033_COMB.seq:US-09-598-982-20
                91:  /cgn2_6/ptodata/2/pna/US6034_COMB.seq:US-09-598-982-20
                92:  /cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-09-598-982-20
                93:  /cgn2_6/ptodata/2/pna/US6036_COMB.seq:US-09-598-982-20
                94:  /cgn2_6/ptodata/2/pna/US6037_COMB.seq:US-09-598-982-20
                95:  /cgn2_6/ptodata/2/pna/US6038_COMB.seq:US-09-598-982-20
                96:  /cgn2_6/ptodata/2/pna/US6039_COMB.seq:US-09-598-982-20
                97:  /cgn2_6/ptodata/2/pna/US6040_COMB.seq:US-09-598-982-20
                98:  /cgn2_6/ptodata/2/pna/US6041_COMB.seq:US-09-598-982-20
                99:  /cgn2_6/ptodata/2/pna/US6042_COMB.seq:US-09-598-982-20
                100:  /cgn2_6/ptodata/2/pna/US6043_COMB.seq:US-09-598-982-20
                101:  /cgn2_6/ptodata/2/pna/US6044_COMB.seq:US-09-598-982-20
                102:  /cgn2_6/ptodata/2/pna/US6045_COMB.seq:US-09-598-982-20
                103:  /cgn2_6/ptodata/2/pna/US6046_COMB.seq:US-09-598-982-20
```

```
         104:  /cgn2_6/ptodata/2/pna/US6047_COMB.seq:US-09-598-982-20
         105:  /cgn2_6/ptodata/2/pna/US6048_COMB.seq:US-09-598-982-20
         106:  /cgn2_6/ptodata/2/pna/US6049_COMB.seq:US-09-598-982-20
         107:  /cgn2_6/ptodata/2/pna/US6050_COMB.seq:US-09-598-982-20
         108:  /cgn2_6/ptodata/2/pna/US6051_COMB.seq:US-09-598-982-20
         109:  /cgn2_6/ptodata/2/pna/US6052_COMB.seq:US-09-598-982-20
         110:  /cgn2_6/ptodata/2/pna/US6053_COMB.seq:US-09-598-982-20

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                                   SUMMARIES
                    %
Result            Query
   No.   Score    Match Length DB  ID                        Description
 -----------------------------------------------------------------------------
     1    1393    100.0    771  24  US-09-598-982-20          Sequence 20, Appl

                                   ALIGNMENTS

RESULT 1
US-09-598-982-20
; Sequence 20, Application US/09598982
; GENERAL INFORMATION:
;  APPLICANT:  NILES, ANDREW L
;  APPLICANT:  MAFFITT, MARK A
;  APPLICANT:  HAAK-FRENDSCHO, MARY
;  TITLE OF INVENTION:  RECOMBINANT PROTEOLYTIC TYPTASES, ACTIVE SITE MUTANTS
;  TITLE OF INVENTION:  THEREOF, AND METHODS OF MAKING SAME
;  FILE REFERENCE:  CIP TRYPTASE
;  CURRENT APPLICATION NUMBER: US/09/598,982
;  CURRENT FILING DATE:  2000-06-21
;  PRIOR APPLICATION NUMBER:  09/079,970
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  51
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 20
;   LENGTH: 771
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (7)..(753)
US-09-598-982-20

Alignment Scores:
Pred. No.:                 0              Length:          771
Score:                     1393.00        Matches:         249
Percent Similarity:        100.00%        Conservative:    0
Best Local Similarity:     100.00%        Mismatches:      0
Query Match:               100.00%        Indels:          0
DB:                        24             Gaps:            0

US-09-598-982-21 (1-249) x US-09-598-982-20 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246
```

```
Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

Search completed: July 23, 2004, 08:27:04
Job time : 1 secs
```

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:          July 23, 2004, 08:28:04 ; Search time 1 Seconds
                                           (without alignments)
                                           0.384 Million cell updates/sec

Title:           US-09-598-982-23
Perfect score:   1395
Sequence:        1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        1 seqs, 771 residues

Total number of hits satisfying chosen parameters:      2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=Pending_Patents_AA_Main:US-09-598-982-23
-DB=Pending_Patents_NA_Main:US-09-598-982-22 -SUFFIX=pto -OUT=align23_22
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       Pending_Patents_NA_Main:US-09-598-982-22
                 1:  /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:US-09-598-982-22
                 2:  /cgn2_6/ptodata/2/pna/US06_COMB.seq:US-09-598-982-22
                 3:  /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-09-598-982-22
                 4:  /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-09-598-982-22
                 5:  /cgn2_6/ptodata/2/pna/US081_COMB.seq:US-09-598-982-22
                 6:  /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-09-598-982-22
                 7:  /cgn2_6/ptodata/2/pna/US083_COMB.seq:US-09-598-982-22
                 8:  /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-09-598-982-22
                 9:  /cgn2_6/ptodata/2/pna/US085_COMB.seq:US-09-598-982-22
                 10:  /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-09-598-982-22
                 11:  /cgn2_6/ptodata/2/pna/US087_COMB.seq:US-09-598-982-22
                 12:  /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-09-598-982-22
                 13:  /cgn2_6/ptodata/2/pna/US089_COMB.seq:US-09-598-982-22
                 14:  /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-598-982-22
                 15:  /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-598-982-22
                 16:  /cgn2_6/ptodata/2/pna/US092A_COMB.seq:US-09-598-982-22
                 17:  /cgn2_6/ptodata/2/pna/US092B_COMB.seq:US-09-598-982-22
                 18:  /cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-598-982-22
                 19:  /cgn2_6/ptodata/2/pna/US093B_COMB.seq:US-09-598-982-22
                 20:  /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-598-982-22
                 21:  /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-598-982-22
                 22:  /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-598-982-22
                 23:  /cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-598-982-22
                 24:  /cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-598-982-22
                 25:  /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-598-982-22
                 26:  /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-598-982-22
                 27:  /cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-598-982-22
                 28:  /cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-598-982-22
                 29:  /cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-598-982-22
                 30:  /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-598-982-22
                 31:  /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-598-982-22
                 32:  /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-598-982-22
                 33:  /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-598-982-22
                 34:  /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-598-982-22
                 35:  /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-598-982-22
                 36:  /cgn2_6/ptodata/2/pna/US098D_COMB.seq:US-09-598-982-22
                 37:  /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-598-982-22
                 38:  /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-598-982-22
                 39:  /cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-598-982-22
                 40:  /cgn2_6/ptodata/2/pna/US099D_COMB.seq:US-09-598-982-22
                 41:  /cgn2_6/ptodata/2/pna/US099E_COMB.seq:US-09-598-982-22
                 42:  /cgn2_6/ptodata/2/pna/US099F_COMB.seq:US-09-598-982-22
                 43:  /cgn2_6/ptodata/2/pna/US100A_COMB.seq:US-09-598-982-22
                 44:  /cgn2_6/ptodata/2/pna/US100B_COMB.seq:US-09-598-982-22
                 45:  /cgn2_6/ptodata/2/pna/US101A_COMB.seq:US-09-598-982-22
                 46:  /cgn2_6/ptodata/2/pna/US101B_COMB.seq:US-09-598-982-22
                 47:  /cgn2_6/ptodata/2/pna/US102A_COMB.seq:US-09-598-982-22
                 48:  /cgn2_6/ptodata/2/pna/US102B_COMB.seq:US-09-598-982-22
                 49:  /cgn2_6/ptodata/2/pna/US103A_COMB.seq:US-09-598-982-22
                 50:  /cgn2_6/ptodata/2/pna/US103B_COMB.seq:US-09-598-982-22
                 51:  /cgn2_6/ptodata/2/pna/US104A_COMB.seq:US-09-598-982-22
                 52:  /cgn2_6/ptodata/2/pna/US104B_COMB.seq:US-09-598-982-22
                 53:  /cgn2_6/ptodata/2/pna/US106_COMB.seq:US-09-598-982-22
                 54:  /cgn2_6/ptodata/2/pna/US107A_COMB.seq:US-09-598-982-22
                 55:  /cgn2_6/ptodata/2/pna/US107B_COMB.seq:US-09-598-982-22
                 56:  /cgn2_6/ptodata/2/pna/US6000_COMB.seq:US-09-598-982-22
                 57:  /cgn2_6/ptodata/2/pna/US6001_COMB.seq:US-09-598-982-22
                 58:  /cgn2_6/ptodata/2/pna/US6002_COMB.seq:US-09-598-982-22
                 59:  /cgn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-598-982-22
                 60:  /cgn2_6/ptodata/2/pna/US6004_COMB.seq:US-09-598-982-22
                 61:  /cgn2_6/ptodata/2/pna/US6005_COMB.seq:US-09-598-982-22
                 62:  /cgn2_6/ptodata/2/pna/US6006_COMB.seq:US-09-598-982-22
                 63:  /cgn2_6/ptodata/2/pna/US6007_COMB.seq:US-09-598-982-22
                 64:  /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-09-598-982-22
                 65:  /cgn2_6/ptodata/2/pna/US6009_COMB.seq:US-09-598-982-22
                 66:  /cgn2_6/ptodata/2/pna/US6010_COMB.seq:US-09-598-982-22
                 67:  /cgn2_6/ptodata/2/pna/US6011_COMB.seq:US-09-598-982-22
                 68:  /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-09-598-982-22
                 69:  /cgn2_6/ptodata/2/pna/US6013_COMB.seq:US-09-598-982-22
                 70:  /cgn2_6/ptodata/2/pna/US6014_COMB.seq:US-09-598-982-22
                 71:  /cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-09-598-982-22
                 72:  /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-09-598-982-22
                 73:  /cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-09-598-982-22
                 74:  /cgn2_6/ptodata/2/pna/US6018_COMB.seq:US-09-598-982-22
                 75:  /cgn2_6/ptodata/2/pna/US6019_COMB.seq:US-09-598-982-22
                 76:  /cgn2_6/ptodata/2/pna/US6020_COMB.seq:US-09-598-982-22
                 77:  /cgn2_6/ptodata/2/pna/US6021_COMB.seq:US-09-598-982-22
                 78:  /cgn2_6/ptodata/2/pna/US6022_COMB.seq:US-09-598-982-22
                 79:  /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:US-09-598-982-22
                 80:  /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:US-09-598-982-22
                 81:  /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-09-598-982-22
                 82:  /cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-09-598-982-22
                 83:  /cgn2_6/ptodata/2/pna/US6026_COMB.seq:US-09-598-982-22
                 84:  /cgn2_6/ptodata/2/pna/US6027_COMB.seq:US-09-598-982-22
                 85:  /cgn2_6/ptodata/2/pna/US6028_COMB.seq:US-09-598-982-22
                 86:  /cgn2_6/ptodata/2/pna/US6029_COMB.seq:US-09-598-982-22
                 87:  /cgn2_6/ptodata/2/pna/US6030_COMB.seq:US-09-598-982-22
                 88:  /cgn2_6/ptodata/2/pna/US6031_COMB.seq:US-09-598-982-22
                 89:  /cgn2_6/ptodata/2/pna/US6032_COMB.seq:US-09-598-982-22
                 90:  /cgn2_6/ptodata/2/pna/US6033_COMB.seq:US-09-598-982-22
                 91:  /cgn2_6/ptodata/2/pna/US6034_COMB.seq:US-09-598-982-22
                 92:  /cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-09-598-982-22
                 93:  /cgn2_6/ptodata/2/pna/US6036_COMB.seq:US-09-598-982-22
                 94:  /cgn2_6/ptodata/2/pna/US6037_COMB.seq:US-09-598-982-22
                 95:  /cgn2_6/ptodata/2/pna/US6038_COMB.seq:US-09-598-982-22
                 96:  /cgn2_6/ptodata/2/pna/US6039_COMB.seq:US-09-598-982-22
                 97:  /cgn2_6/ptodata/2/pna/US6040_COMB.seq:US-09-598-982-22
                 98:  /cgn2_6/ptodata/2/pna/US6041_COMB.seq:US-09-598-982-22
                 99:  /cgn2_6/ptodata/2/pna/US6042_COMB.seq:US-09-598-982-22
                 100:  /cgn2_6/ptodata/2/pna/US6043_COMB.seq:US-09-598-982-22
                 101:  /cgn2_6/ptodata/2/pna/US6044_COMB.seq:US-09-598-982-22
                 102:  /cgn2_6/ptodata/2/pna/US6045_COMB.seq:US-09-598-982-22
                 103:  /cgn2_6/ptodata/2/pna/US6046_COMB.seq:US-09-598-982-22
```

```
          104:  /cgn2_6/ptodata/2/pna/US6047_COMB.seq:US-09-598-982-22
          105:  /cgn2_6/ptodata/2/pna/US6048_COMB.seq:US-09-598-982-22
          106:  /cgn2_6/ptodata/2/pna/US6049_COMB.seq:US-09-598-982-22
          107:  /cgn2_6/ptodata/2/pna/US6050_COMB.seq:US-09-598-982-22
          108:  /cgn2_6/ptodata/2/pna/US6051_COMB.seq:US-09-598-982-22
          109:  /cgn2_6/ptodata/2/pna/US6052_COMB.seq:US-09-598-982-22
          110:  /cgn2_6/ptodata/2/pna/US6053_COMB.seq:US-09-598-982-22

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                   SUMMARIES
                     %
Result             Query
  No.     Score    Match Length DB  ID                        Description
-------------------------------------------------------------------------------
    1     1395   100.0    771  24  US-09-598-982-22            Sequence 22, Appl

                                   ALIGNMENTS

RESULT 1
US-09-598-982-22
; Sequence 22, Application US/09598982
; GENERAL INFORMATION:
;  APPLICANT:  NILES, ANDREW L
;  APPLICANT:  MAFFITT, MARK A
;  APPLICANT:  HAAK-FRENDSCHO, MARY
;  TITLE OF INVENTION:  RECOMBINANT PROTEOLYTIC TYPTASES, ACTIVE SITE MUTANTS
;  TITLE OF INVENTION:  THEREOF, AND METHODS OF MAKING SAME
;  FILE REFERENCE:  CIP TRYPTASE
;  CURRENT APPLICATION NUMBER: US/09/598,982
;  CURRENT FILING DATE:  2000-06-21
;  PRIOR APPLICATION NUMBER:  09/079,970
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  51
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 22
;   LENGTH: 771
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (7)..(753)
US-09-598-982-22

Alignment Scores:
Pred. No.:                0            Length:         771
Score:                    1395.00      Matches:        249
Percent Similarity:       100.00%      Conservative:   0
Best Local Similarity:    100.00%      Mismatches:     0
Query Match:              100.00%      Indels:         0
DB:                       24           Gaps:           0

US-09-598-982-23 (1-249) x US-09-598-982-22 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246
```

```
Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAlaIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

Search completed: July 23, 2004, 08:28:05
Job time : 1 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 23, 2004, 08:29:03 ; Search time 0.001 Seconds
                                          (without alignments)
                                          383.958 Million cell updates/sec

Title:          US-09-598-982-25
Perfect score:  1397
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1 seqs, 771 residues

Total number of hits satisfying chosen parameters:      2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=Pending_Patents_AA_Main:US-09-598-982-25
-DB=Pending_Patents_NA_Main:US-09-598-982-24 -SUFFIX=pto -OUT=align25_24
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Pending_Patents_NA_Main:US-09-598-982-24
                1:  /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:US-09-598-982-24
                2:  /cgn2_6/ptodata/2/pna/US06_COMB.seq:US-09-598-982-24
                3:  /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-09-598-982-24
                4:  /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-09-598-982-24
                5:  /cgn2_6/ptodata/2/pna/US081_COMB.seq:US-09-598-982-24
                6:  /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-09-598-982-24
                7:  /cgn2_6/ptodata/2/pna/US083_COMB.seq:US-09-598-982-24
                8:  /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-09-598-982-24
                9:  /cgn2_6/ptodata/2/pna/US085_COMB.seq:US-09-598-982-24
                10:  /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-09-598-982-24
                11:  /cgn2_6/ptodata/2/pna/US087_COMB.seq:US-09-598-982-24
                12:  /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-09-598-982-24
                13:  /cgn2_6/ptodata/2/pna/US089_COMB.seq:US-09-598-982-24
                14:  /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-598-982-24
                15:  /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-598-982-24
                16:  /cgn2_6/ptodata/2/pna/US092A_COMB.seq:US-09-598-982-24
                17:  /cgn2_6/ptodata/2/pna/US092B_COMB.seq:US-09-598-982-24
                18:  /cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-598-982-24
                19:  /cgn2_6/ptodata/2/pna/US093B_COMB.seq:US-09-598-982-24
                20:  /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-598-982-24
                21:  /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-598-982-24
                22:  /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-598-982-24
                23:  /cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-598-982-24
                24:  /cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-598-982-24
                25:  /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-598-982-24
                26:  /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-598-982-24
                27:  /cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-598-982-24
                28:  /cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-598-982-24
                29:  /cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-598-982-24
                30:  /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-598-982-24
```

```
                31:  /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-598-982-24
                32:  /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-598-982-24
                33:  /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-598-982-24
                34:  /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-598-982-24
                35:  /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-598-982-24
                36:  /cgn2_6/ptodata/2/pna/US098D_COMB.seq:US-09-598-982-24
                37:  /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-598-982-24
                38:  /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-598-982-24
                39:  /cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-598-982-24
                40:  /cgn2_6/ptodata/2/pna/US099D_COMB.seq:US-09-598-982-24
                41:  /cgn2_6/ptodata/2/pna/US099E_COMB.seq:US-09-598-982-24
                42:  /cgn2_6/ptodata/2/pna/US099F_COMB.seq:US-09-598-982-24
                43:  /cgn2_6/ptodata/2/pna/US100A_COMB.seq:US-09-598-982-24
                44:  /cgn2_6/ptodata/2/pna/US100B_COMB.seq:US-09-598-982-24
                45:  /cgn2_6/ptodata/2/pna/US101A_COMB.seq:US-09-598-982-24
                46:  /cgn2_6/ptodata/2/pna/US101B_COMB.seq:US-09-598-982-24
                47:  /cgn2_6/ptodata/2/pna/US102A_COMB.seq:US-09-598-982-24
                48:  /cgn2_6/ptodata/2/pna/US102B_COMB.seq:US-09-598-982-24
                49:  /cgn2_6/ptodata/2/pna/US103A_COMB.seq:US-09-598-982-24
                50:  /cgn2_6/ptodata/2/pna/US103B_COMB.seq:US-09-598-982-24
                51:  /cgn2_6/ptodata/2/pna/US104A_COMB.seq:US-09-598-982-24
                52:  /cgn2_6/ptodata/2/pna/US104B_COMB.seq:US-09-598-982-24
                53:  /cgn2_6/ptodata/2/pna/US106_COMB.seq:US-09-598-982-24
                54:  /cgn2_6/ptodata/2/pna/US107A_COMB.seq:US-09-598-982-24
                55:  /cgn2_6/ptodata/2/pna/US107B_COMB.seq:US-09-598-982-24
                56:  /cgn2_6/ptodata/2/pna/US6000_COMB.seq:US-09-598-982-24
                57:  /cgn2_6/ptodata/2/pna/US6001_COMB.seq:US-09-598-982-24
                58:  /cgn2_6/ptodata/2/pna/US6002_COMB.seq:US-09-598-982-24
                59:  /cgn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-598-982-24
                60:  /cgn2_6/ptodata/2/pna/US6004_COMB.seq:US-09-598-982-24
                61:  /cgn2_6/ptodata/2/pna/US6005_COMB.seq:US-09-598-982-24
                62:  /cgn2_6/ptodata/2/pna/US6006_COMB.seq:US-09-598-982-24
                63:  /cgn2_6/ptodata/2/pna/US6007_COMB.seq:US-09-598-982-24
                64:  /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-09-598-982-24
                65:  /cgn2_6/ptodata/2/pna/US6009_COMB.seq:US-09-598-982-24
                66:  /cgn2_6/ptodata/2/pna/US6010_COMB.seq:US-09-598-982-24
                67:  /cgn2_6/ptodata/2/pna/US6011_COMB.seq:US-09-598-982-24
                68:  /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-09-598-982-24
                69:  /cgn2_6/ptodata/2/pna/US6013_COMB.seq:US-09-598-982-24
                70:  /cgn2_6/ptodata/2/pna/US6014_COMB.seq:US-09-598-982-24
                71:  /cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-09-598-982-24
                72:  /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-09-598-982-24
                73:  /cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-09-598-982-24
                74:  /cgn2_6/ptodata/2/pna/US6018_COMB.seq:US-09-598-982-24
                75:  /cgn2_6/ptodata/2/pna/US6019_COMB.seq:US-09-598-982-24
                76:  /cgn2_6/ptodata/2/pna/US6020_COMB.seq:US-09-598-982-24
                77:  /cgn2_6/ptodata/2/pna/US6021_COMB.seq:US-09-598-982-24
                78:  /cgn2_6/ptodata/2/pna/US6022_COMB.seq:US-09-598-982-24
                79:  /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:US-09-598-982-24
                80:  /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:US-09-598-982-24
                81:  /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-09-598-982-24
                82:  /cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-09-598-982-24
                83:  /cgn2_6/ptodata/2/pna/US6026_COMB.seq:US-09-598-982-24
                84:  /cgn2_6/ptodata/2/pna/US6027_COMB.seq:US-09-598-982-24
                85:  /cgn2_6/ptodata/2/pna/US6028_COMB.seq:US-09-598-982-24
                86:  /cgn2_6/ptodata/2/pna/US6029_COMB.seq:US-09-598-982-24
                87:  /cgn2_6/ptodata/2/pna/US6030_COMB.seq:US-09-598-982-24
                88:  /cgn2_6/ptodata/2/pna/US6031_COMB.seq:US-09-598-982-24
                89:  /cgn2_6/ptodata/2/pna/US6032_COMB.seq:US-09-598-982-24
                90:  /cgn2_6/ptodata/2/pna/US6033_COMB.seq:US-09-598-982-24
                91:  /cgn2_6/ptodata/2/pna/US6034_COMB.seq:US-09-598-982-24
                92:  /cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-09-598-982-24
                93:  /cgn2_6/ptodata/2/pna/US6036_COMB.seq:US-09-598-982-24
                94:  /cgn2_6/ptodata/2/pna/US6037_COMB.seq:US-09-598-982-24
                95:  /cgn2_6/ptodata/2/pna/US6038_COMB.seq:US-09-598-982-24
                96:  /cgn2_6/ptodata/2/pna/US6039_COMB.seq:US-09-598-982-24
                97:  /cgn2_6/ptodata/2/pna/US6040_COMB.seq:US-09-598-982-24
                98:  /cgn2_6/ptodata/2/pna/US6041_COMB.seq:US-09-598-982-24
                99:  /cgn2_6/ptodata/2/pna/US6042_COMB.seq:US-09-598-982-24
                100:  /cgn2_6/ptodata/2/pna/US6043_COMB.seq:US-09-598-982-24
                101:  /cgn2_6/ptodata/2/pna/US6044_COMB.seq:US-09-598-982-24
                102:  /cgn2_6/ptodata/2/pna/US6045_COMB.seq:US-09-598-982-24
                103:  /cgn2_6/ptodata/2/pna/US6046_COMB.seq:US-09-598-982-24
```

```
          104:  /cgn2_6/ptodata/2/pna/US6047_COMB.seq:US-09-598-982-24
          105:  /cgn2_6/ptodata/2/pna/US6048_COMB.seq:US-09-598-982-24
          106:  /cgn2_6/ptodata/2/pna/US6049_COMB.seq:US-09-598-982-24
          107:  /cgn2_6/ptodata/2/pna/US6050_COMB.seq:US-09-598-982-24
          108:  /cgn2_6/ptodata/2/pna/US6051_COMB.seq:US-09-598-982-24
          109:  /cgn2_6/ptodata/2/pna/US6052_COMB.seq:US-09-598-982-24
          110:  /cgn2_6/ptodata/2/pna/US6053_COMB.seq:US-09-598-982-24

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                   SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                          Description
 ---------------------------------------------------------------------------
    1    1397   100.0    771  24  US-09-598-982-24             Sequence 24, Appl

                                   ALIGNMENTS

RESULT 1
US-09-598-982-24
; Sequence 24, Application US/09598982
; GENERAL INFORMATION:
;  APPLICANT:  NILES, ANDREW L
;  APPLICANT:  MAFFITT, MARK A
;  APPLICANT:  HAAK-FRENDSCHO, MARY
;  TITLE OF INVENTION:  RECOMBINANT PROTEOLYTIC TYPTASES, ACTIVE SITE MUTANTS
;  TITLE OF INVENTION:  THEREOF, AND METHODS OF MAKING SAME
;  FILE REFERENCE:  CIP TRYPTASE
;  CURRENT APPLICATION NUMBER: US/09/598,982
;  CURRENT FILING DATE:  2000-06-21
;  PRIOR APPLICATION NUMBER:  09/079,970
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  51
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 24
;   LENGTH: 771
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (7)..(753)
US-09-598-982-24

Alignment Scores:
Pred. No.:               0              Length:         771
Score:                   1397.00        Matches:        249
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             100.00%        Indels:         0
DB:                      24             Gaps:           0

US-09-598-982-25 (1-249) x US-09-598-982-24 (1-771)

Qy           1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy          21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy          41 GlnTrpValLeuThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeu 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy          61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy          81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy         101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy         121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy         141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy         161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy         181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspAlaGlyGly 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCCGGCGGA 606

Qy         201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         607 CCTCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy         221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy         241 IleHisHisTyrValProLysLysPro 249
               |||||||||||||||||||||||||||
Db         727 ATCCACCACTATGTCCCCAAAAAGCCG 753

Search completed: July 23, 2004, 08:29:04
Job time : 0.001 secs
```

```
                         GenCore version 5.1.6
                Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 23, 2004, 08:29:57 ; Search time 1 Seconds
                                           (without alignments)
                                          0.384 Million cell updates/sec

Title:          US-09-598-982-27
Perfect score:  1397
Sequence:       1 LEKRIVGGQEAPRSKWPWQV.........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1 seqs, 771 residues

Total number of hits satisfying chosen parameters:      2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=Pending_Patents_AA_Main:US-09-598-982-27
-DB=Pending_Patents_NA_Main:US-09-598-982-26 -SUFFIX=pto -OUT=align27_26
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Pending_Patents_NA_Main:US-09-598-982-26
                1:  /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:US-09-598-982-26
                2:  /cgn2_6/ptodata/2/pna/US06_COMB.seq:US-09-598-982-26
                3:  /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-09-598-982-26
                4:  /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-09-598-982-26
                5:  /cgn2_6/ptodata/2/pna/US081_COMB.seq:US-09-598-982-26
                6:  /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-09-598-982-26
                7:  /cgn2_6/ptodata/2/pna/US083_COMB.seq:US-09-598-982-26
                8:  /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-09-598-982-26
                9:  /cgn2_6/ptodata/2/pna/US085_COMB.seq:US-09-598-982-26
                10:  /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-09-598-982-26
                11:  /cgn2_6/ptodata/2/pna/US087_COMB.seq:US-09-598-982-26
                12:  /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-09-598-982-26
                13:  /cgn2_6/ptodata/2/pna/US089_COMB.seq:US-09-598-982-26
                14:  /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-598-982-26
                15:  /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-598-982-26
                16:  /cgn2_6/ptodata/2/pna/US092A_COMB.seq:US-09-598-982-26
                17:  /cgn2_6/ptodata/2/pna/US092B_COMB.seq:US-09-598-982-26
                18:  /cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-598-982-26
                19:  /cgn2_6/ptodata/2/pna/US093B_COMB.seq:US-09-598-982-26
                20:  /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-598-982-26
                21:  /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-598-982-26
                22:  /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-598-982-26
                23:  /cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-598-982-26
                24:  /cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-598-982-26
                25:  /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-598-982-26
                26:  /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-598-982-26
                27:  /cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-598-982-26
                28:  /cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-598-982-26
                29:  /cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-598-982-26
                30:  /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-598-982-26
                31:  /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-598-982-26
                32:  /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-598-982-26
                33:  /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-598-982-26
                34:  /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-598-982-26
                35:  /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-598-982-26
                36:  /cgn2_6/ptodata/2/pna/US098D_COMB.seq:US-09-598-982-26
                37:  /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-598-982-26
                38:  /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-598-982-26
                39:  /cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-598-982-26
                40:  /cgn2_6/ptodata/2/pna/US099D_COMB.seq:US-09-598-982-26
                41:  /cgn2_6/ptodata/2/pna/US099E_COMB.seq:US-09-598-982-26
                42:  /cgn2_6/ptodata/2/pna/US099F_COMB.seq:US-09-598-982-26
                43:  /cgn2_6/ptodata/2/pna/US100A_COMB.seq:US-09-598-982-26
                44:  /cgn2_6/ptodata/2/pna/US100B_COMB.seq:US-09-598-982-26
                45:  /cgn2_6/ptodata/2/pna/US101A_COMB.seq:US-09-598-982-26
                46:  /cgn2_6/ptodata/2/pna/US101B_COMB.seq:US-09-598-982-26
                47:  /cgn2_6/ptodata/2/pna/US102A_COMB.seq:US-09-598-982-26
                48:  /cgn2_6/ptodata/2/pna/US102B_COMB.seq:US-09-598-982-26
                49:  /cgn2_6/ptodata/2/pna/US103A_COMB.seq:US-09-598-982-26
                50:  /cgn2_6/ptodata/2/pna/US103B_COMB.seq:US-09-598-982-26
                51:  /cgn2_6/ptodata/2/pna/US104A_COMB.seq:US-09-598-982-26
                52:  /cgn2_6/ptodata/2/pna/US104B_COMB.seq:US-09-598-982-26
                53:  /cgn2_6/ptodata/2/pna/US106_COMB.seq:US-09-598-982-26
                54:  /cgn2_6/ptodata/2/pna/US107A_COMB.seq:US-09-598-982-26
                55:  /cgn2_6/ptodata/2/pna/US107B_COMB.seq:US-09-598-982-26
                56:  /cgn2_6/ptodata/2/pna/US6000_COMB.seq:US-09-598-982-26
                57:  /cgn2_6/ptodata/2/pna/US6001_COMB.seq:US-09-598-982-26
                58:  /cgn2_6/ptodata/2/pna/US6002_COMB.seq:US-09-598-982-26
                59:  /cgn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-598-982-26
                60:  /cgn2_6/ptodata/2/pna/US6004_COMB.seq:US-09-598-982-26
                61:  /cgn2_6/ptodata/2/pna/US6005_COMB.seq:US-09-598-982-26
                62:  /cgn2_6/ptodata/2/pna/US6006_COMB.seq:US-09-598-982-26
                63:  /cgn2_6/ptodata/2/pna/US6007_COMB.seq:US-09-598-982-26
                64:  /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-09-598-982-26
                65:  /cgn2_6/ptodata/2/pna/US6009_COMB.seq:US-09-598-982-26
                66:  /cgn2_6/ptodata/2/pna/US6010_COMB.seq:US-09-598-982-26
                67:  /cgn2_6/ptodata/2/pna/US6011_COMB.seq:US-09-598-982-26
                68:  /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-09-598-982-26
                69:  /cgn2_6/ptodata/2/pna/US6013_COMB.seq:US-09-598-982-26
                70:  /cgn2_6/ptodata/2/pna/US6014_COMB.seq:US-09-598-982-26
                71:  /cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-09-598-982-26
                72:  /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-09-598-982-26
                73:  /cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-09-598-982-26
                74:  /cgn2_6/ptodata/2/pna/US6018_COMB.seq:US-09-598-982-26
                75:  /cgn2_6/ptodata/2/pna/US6019_COMB.seq:US-09-598-982-26
                76:  /cgn2_6/ptodata/2/pna/US6020_COMB.seq:US-09-598-982-26
                77:  /cgn2_6/ptodata/2/pna/US6021_COMB.seq:US-09-598-982-26
                78:  /cgn2_6/ptodata/2/pna/US6022_COMB.seq:US-09-598-982-26
                79:  /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:US-09-598-982-26
                80:  /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:US-09-598-982-26
                81:  /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-09-598-982-26
                82:  /cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-09-598-982-26
                83:  /cgn2_6/ptodata/2/pna/US6026_COMB.seq:US-09-598-982-26
                84:  /cgn2_6/ptodata/2/pna/US6027_COMB.seq:US-09-598-982-26
                85:  /cgn2_6/ptodata/2/pna/US6028_COMB.seq:US-09-598-982-26
                86:  /cgn2_6/ptodata/2/pna/US6029_COMB.seq:US-09-598-982-26
                87:  /cgn2_6/ptodata/2/pna/US6030_COMB.seq:US-09-598-982-26
                88:  /cgn2_6/ptodata/2/pna/US6031_COMB.seq:US-09-598-982-26
                89:  /cgn2_6/ptodata/2/pna/US6032_COMB.seq:US-09-598-982-26
                90:  /cgn2_6/ptodata/2/pna/US6033_COMB.seq:US-09-598-982-26
                91:  /cgn2_6/ptodata/2/pna/US6034_COMB.seq:US-09-598-982-26
                92:  /cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-09-598-982-26
                93:  /cgn2_6/ptodata/2/pna/US6036_COMB.seq:US-09-598-982-26
                94:  /cgn2_6/ptodata/2/pna/US6037_COMB.seq:US-09-598-982-26
                95:  /cgn2_6/ptodata/2/pna/US6038_COMB.seq:US-09-598-982-26
                96:  /cgn2_6/ptodata/2/pna/US6039_COMB.seq:US-09-598-982-26
                97:  /cgn2_6/ptodata/2/pna/US6040_COMB.seq:US-09-598-982-26
                98:  /cgn2_6/ptodata/2/pna/US6041_COMB.seq:US-09-598-982-26
                99:  /cgn2_6/ptodata/2/pna/US6042_COMB.seq:US-09-598-982-26
                100:  /cgn2_6/ptodata/2/pna/US6043_COMB.seq:US-09-598-982-26
                101:  /cgn2_6/ptodata/2/pna/US6044_COMB.seq:US-09-598-982-26
                102:  /cgn2_6/ptodata/2/pna/US6045_COMB.seq:US-09-598-982-26
                103:  /cgn2_6/ptodata/2/pna/US6046_COMB.seq:US-09-598-982-26
```

```
         104:  /cgn2_6/ptodata/2/pna/US6047_COMB.seq:US-09-598-982-26
         105:  /cgn2_6/ptodata/2/pna/US6048_COMB.seq:US-09-598-982-26
         106:  /cgn2_6/ptodata/2/pna/US6049_COMB.seq:US-09-598-982-26
         107:  /cgn2_6/ptodata/2/pna/US6050_COMB.seq:US-09-598-982-26
         108:  /cgn2_6/ptodata/2/pna/US6051_COMB.seq:US-09-598-982-26
         109:  /cgn2_6/ptodata/2/pna/US6052_COMB.seq:US-09-598-982-26
         110:  /cgn2_6/ptodata/2/pna/US6053_COMB.seq:US-09-598-982-26

   Pred. No. is the number of results predicted by chance to have a
   score greater than or equal to the score of the result being printed,
   and is derived by analysis of the total score distribution.

                               SUMMARIES
                    %
Result            Query
  No.    Score    Match Length DB  ID                          Description
 ---------------------------------------------------------------------------
    1     1397    100.0    771  24  US-09-598-982-26            Sequence 26, Appl


                               ALIGNMENTS

RESULT 1
US-09-598-982-26
; Sequence 26, Application US/09598982
; GENERAL INFORMATION:
;  APPLICANT:  NILES, ANDREW L
;  APPLICANT:  MAFFITT, MARK A
;  APPLICANT:  HAAK-FRENDSCHO, MARY
;  TITLE OF INVENTION:  RECOMBINANT PROTEOLYTIC TYPTASES, ACTIVE SITE MUTANTS
;  TITLE OF INVENTION:  THEREOF, AND METHODS OF MAKING SAME
;  FILE REFERENCE:  CIP TRYPTASE
;  CURRENT APPLICATION NUMBER: US/09/598,982
;  CURRENT FILING DATE:  2000-06-21
;  PRIOR APPLICATION NUMBER:  09/079,970
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  51
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 26
;   LENGTH: 771
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (7)..(753)
US-09-598-982-26

Alignment Scores:
Pred. No.:                0               Length:         771
Score:                    1397.00         Matches:        249
Percent Similarity:       100.00%         Conservative:   0
Best Local Similarity:    100.00%         Mismatches:     0
Query Match:              100.00%         Indels:         0
DB:                       24              Gaps:           0

US-09-598-982-27 (1-249) x US-09-598-982-26 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246
```

```
Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspAlaGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCCGGCGGA 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCACTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

Search completed: July 23, 2004, 08:29:58
Job time : 1 secs
```

```
                 GenCore version 5.1.6
         Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 23, 2004, 08:34:09 ; Search time 1 Seconds
                                          (without alignments)
                                          0.384 Million cell updates/sec

Title:          US-09-598-982-37
Perfect score:  1394
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1 seqs, 771 residues

Total number of hits satisfying chosen parameters:      2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=Pending_Patents_AA_Main:US-09-598-982-37
-DB=Pending_Patents_NA_Main:US-09-598-982-36 -SUFFIX=pto -OUT=align37_36
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Pending_Patents_NA_Main:US-09-598-982-36
                1:  /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:US-09-598-982-36
                2:  /cgn2_6/ptodata/2/pna/US06_COMB.seq:US-09-598-982-36
                3:  /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-09-598-982-36
                4:  /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-09-598-982-36
                5:  /cgn2_6/ptodata/2/pna/US081_COMB.seq:US-09-598-982-36
                6:  /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-09-598-982-36
                7:  /cgn2_6/ptodata/2/pna/US083_COMB.seq:US-09-598-982-36
                8:  /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-09-598-982-36
                9:  /cgn2_6/ptodata/2/pna/US085_COMB.seq:US-09-598-982-36
                10:  /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-09-598-982-36
                11:  /cgn2_6/ptodata/2/pna/US087_COMB.seq:US-09-598-982-36
                12:  /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-09-598-982-36
                13:  /cgn2_6/ptodata/2/pna/US089_COMB.seq:US-09-598-982-36
                14:  /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-598-982-36
                15:  /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-598-982-36
                16:  /cgn2_6/ptodata/2/pna/US092A_COMB.seq:US-09-598-982-36
                17:  /cgn2_6/ptodata/2/pna/US092B_COMB.seq:US-09-598-982-36
                18:  /cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-598-982-36
                19:  /cgn2_6/ptodata/2/pna/US093B_COMB.seq:US-09-598-982-36
                20:  /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-598-982-36
                21:  /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-598-982-36
                22:  /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-598-982-36
                23:  /cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-598-982-36
                24:  /cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-598-982-36
                25:  /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-598-982-36
                26:  /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-598-982-36
                27:  /cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-598-982-36
                28:  /cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-598-982-36
                29:  /cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-598-982-36
                30:  /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-598-982-36
                31:  /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-598-982-36
                32:  /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-598-982-36
                33:  /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-598-982-36
                34:  /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-598-982-36
                35:  /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-598-982-36
                36:  /cgn2_6/ptodata/2/pna/US098D_COMB.seq:US-09-598-982-36
                37:  /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-598-982-36
                38:  /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-598-982-36
                39:  /cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-598-982-36
                40:  /cgn2_6/ptodata/2/pna/US099D_COMB.seq:US-09-598-982-36
                41:  /cgn2_6/ptodata/2/pna/US099E_COMB.seq:US-09-598-982-36
                42:  /cgn2_6/ptodata/2/pna/US099F_COMB.seq:US-09-598-982-36
                43:  /cgn2_6/ptodata/2/pna/US100A_COMB.seq:US-09-598-982-36
                44:  /cgn2_6/ptodata/2/pna/US100B_COMB.seq:US-09-598-982-36
                45:  /cgn2_6/ptodata/2/pna/US101A_COMB.seq:US-09-598-982-36
                46:  /cgn2_6/ptodata/2/pna/US101B_COMB.seq:US-09-598-982-36
                47:  /cgn2_6/ptodata/2/pna/US102A_COMB.seq:US-09-598-982-36
                48:  /cgn2_6/ptodata/2/pna/US102B_COMB.seq:US-09-598-982-36
                49:  /cgn2_6/ptodata/2/pna/US103A_COMB.seq:US-09-598-982-36
                50:  /cgn2_6/ptodata/2/pna/US103B_COMB.seq:US-09-598-982-36
                51:  /cgn2_6/ptodata/2/pna/US104A_COMB.seq:US-09-598-982-36
                52:  /cgn2_6/ptodata/2/pna/US104B_COMB.seq:US-09-598-982-36
                53:  /cgn2_6/ptodata/2/pna/US106_COMB.seq:US-09-598-982-36
                54:  /cgn2_6/ptodata/2/pna/US107A_COMB.seq:US-09-598-982-36
                55:  /cgn2_6/ptodata/2/pna/US107B_COMB.seq:US-09-598-982-36
                56:  /cgn2_6/ptodata/2/pna/US6000_COMB.seq:US-09-598-982-36
                57:  /cgn2_6/ptodata/2/pna/US6001_COMB.seq:US-09-598-982-36
                58:  /cgn2_6/ptodata/2/pna/US6002_COMB.seq:US-09-598-982-36
                59:  /cgn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-598-982-36
                60:  /cgn2_6/ptodata/2/pna/US6004_COMB.seq:US-09-598-982-36
                61:  /cgn2_6/ptodata/2/pna/US6005_COMB.seq:US-09-598-982-36
                62:  /cgn2_6/ptodata/2/pna/US6006_COMB.seq:US-09-598-982-36
                63:  /cgn2_6/ptodata/2/pna/US6007_COMB.seq:US-09-598-982-36
                64:  /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-09-598-982-36
                65:  /cgn2_6/ptodata/2/pna/US6009_COMB.seq:US-09-598-982-36
                66:  /cgn2_6/ptodata/2/pna/US6010_COMB.seq:US-09-598-982-36
                67:  /cgn2_6/ptodata/2/pna/US6011_COMB.seq:US-09-598-982-36
                68:  /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-09-598-982-36
                69:  /cgn2_6/ptodata/2/pna/US6013_COMB.seq:US-09-598-982-36
                70:  /cgn2_6/ptodata/2/pna/US6014_COMB.seq:US-09-598-982-36
                71:  /cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-09-598-982-36
                72:  /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-09-598-982-36
                73:  /cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-09-598-982-36
                74:  /cgn2_6/ptodata/2/pna/US6018_COMB.seq:US-09-598-982-36
                75:  /cgn2_6/ptodata/2/pna/US6019_COMB.seq:US-09-598-982-36
                76:  /cgn2_6/ptodata/2/pna/US6020_COMB.seq:US-09-598-982-36
                77:  /cgn2_6/ptodata/2/pna/US6021_COMB.seq:US-09-598-982-36
                78:  /cgn2_6/ptodata/2/pna/US6022_COMB.seq:US-09-598-982-36
                79:  /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:US-09-598-982-36
                80:  /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:US-09-598-982-36
                81:  /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-09-598-982-36
                82:  /cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-09-598-982-36
                83:  /cgn2_6/ptodata/2/pna/US6026_COMB.seq:US-09-598-982-36
                84:  /cgn2_6/ptodata/2/pna/US6027_COMB.seq:US-09-598-982-36
                85:  /cgn2_6/ptodata/2/pna/US6028_COMB.seq:US-09-598-982-36
                86:  /cgn2_6/ptodata/2/pna/US6029_COMB.seq:US-09-598-982-36
                87:  /cgn2_6/ptodata/2/pna/US6030_COMB.seq:US-09-598-982-36
                88:  /cgn2_6/ptodata/2/pna/US6031_COMB.seq:US-09-598-982-36
                89:  /cgn2_6/ptodata/2/pna/US6032_COMB.seq:US-09-598-982-36
                90:  /cgn2_6/ptodata/2/pna/US6033_COMB.seq:US-09-598-982-36
                91:  /cgn2_6/ptodata/2/pna/US6034_COMB.seq:US-09-598-982-36
                92:  /cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-09-598-982-36
                93:  /cgn2_6/ptodata/2/pna/US6036_COMB.seq:US-09-598-982-36
                94:  /cgn2_6/ptodata/2/pna/US6037_COMB.seq:US-09-598-982-36
                95:  /cgn2_6/ptodata/2/pna/US6038_COMB.seq:US-09-598-982-36
                96:  /cgn2_6/ptodata/2/pna/US6039_COMB.seq:US-09-598-982-36
                97:  /cgn2_6/ptodata/2/pna/US6040_COMB.seq:US-09-598-982-36
                98:  /cgn2_6/ptodata/2/pna/US6041_COMB.seq:US-09-598-982-36
                99:  /cgn2_6/ptodata/2/pna/US6042_COMB.seq:US-09-598-982-36
                100:  /cgn2_6/ptodata/2/pna/US6043_COMB.seq:US-09-598-982-36
                101:  /cgn2_6/ptodata/2/pna/US6044_COMB.seq:US-09-598-982-36
                102:  /cgn2_6/ptodata/2/pna/US6045_COMB.seq:US-09-598-982-36
                103:  /cgn2_6/ptodata/2/pna/US6046_COMB.seq:US-09-598-982-36
```

```
          104:  /cgn2_6/ptodata/2/pna/US6047_COMB.seq:US-09-598-982-36
          105:  /cgn2_6/ptodata/2/pna/US6048_COMB.seq:US-09-598-982-36
          106:  /cgn2_6/ptodata/2/pna/US6049_COMB.seq:US-09-598-982-36
          107:  /cgn2_6/ptodata/2/pna/US6050_COMB.seq:US-09-598-982-36
          108:  /cgn2_6/ptodata/2/pna/US6051_COMB.seq:US-09-598-982-36
          109:  /cgn2_6/ptodata/2/pna/US6052_COMB.seq:US-09-598-982-36
          110:  /cgn2_6/ptodata/2/pna/US6053_COMB.seq:US-09-598-982-36

    Pred. No. is the number of results predicted by chance to have a
    score greater than or equal to the score of the result being printed,
    and is derived by analysis of the total score distribution.

                                SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                    Description
-------------------------------------------------------------------------
    1    1394   100.0    771  24  US-09-598-982-36 .        Sequence 36, Appl

                                ALIGNMENTS

RESULT 1
US-09-598-982-36
; Sequence 36, Application US/09598982
; GENERAL INFORMATION:
;  APPLICANT:  NILES, ANDREW L
;  APPLICANT:  MAFFITT, MARK A
;  APPLICANT:  HAAK-FRENDSCHO, MARY
;  TITLE OF INVENTION:  RECOMBINANT PROTEOLYTIC TYPTASES, ACTIVE SITE MUTANTS
;  TITLE OF INVENTION:  THEREOF, AND METHODS OF MAKING SAME
;  FILE REFERENCE:  CIP TRYPTASE
;  CURRENT APPLICATION NUMBER: US/09/598,982
;  CURRENT FILING DATE:  2000-06-21
;  PRIOR APPLICATION NUMBER:  09/079,970
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  51
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 36
;   LENGTH: 771
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (7)..(753)
US-09-598-982-36

Alignment Scores:
Pred. No.:              0              Length:        771
Score:                  1394.00        Matches:       249
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     24             Gaps:          0

US-09-598-982-37 (1-249) x US-09-598-982-36 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246
```

```
Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValAsnValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

Search completed: July 23, 2004, 08:34:10
Job time : 1 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 23, 2004, 08:35:02 ; Search time 0.001 Seconds
                                           (without alignments)
                                           383.958 Million cell updates/sec

Title:          US-09-598-982-39
Perfect score:  1396
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1 seqs, 771 residues

Total number of hits satisfying chosen parameters:      2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=Pending_Patents_AA_Main:US-09-598-982-39
-DB=Pending_Patents_NA_Main:US-09-598-982-38 -SUFFIX=pto -OUT=align39_38
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Pending_Patents_NA_Main:US-09-598-982-38
                1:  /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:US-09-598-982-38
                2:  /cgn2_6/ptodata/2/pna/US06_COMB.seq:US-09-598-982-38
                3:  /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-09-598-982-38
                4:  /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-09-598-982-38
                5:  /cgn2_6/ptodata/2/pna/US081_COMB.seq:US-09-598-982-38
                6:  /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-09-598-982-38
                7:  /cgn2_6/ptodata/2/pna/US083_COMB.seq:US-09-598-982-38
                8:  /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-09-598-982-38
                9:  /cgn2_6/ptodata/2/pna/US085_COMB.seq:US-09-598-982-38
                10:  /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-09-598-982-38
                11:  /cgn2_6/ptodata/2/pna/US087_COMB.seq:US-09-598-982-38
                12:  /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-09-598-982-38
                13:  /cgn2_6/ptodata/2/pna/US089_COMB.seq:US-09-598-982-38
                14:  /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-598-982-38
                15:  /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-598-982-38
                16:  /cgn2_6/ptodata/2/pna/US092A_COMB.seq:US-09-598-982-38
                17:  /cgn2_6/ptodata/2/pna/US092B_COMB.seq:US-09-598-982-38
                18:  /cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-598-982-38
                19:  /cgn2_6/ptodata/2/pna/US093B_COMB.seq:US-09-598-982-38
                20:  /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-598-982-38
                21:  /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-598-982-38
                22:  /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-598-982-38
                23:  /cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-598-982-38
                24:  /cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-598-982-38
                25:  /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-598-982-38
                26:  /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-598-982-38
                27:  /cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-598-982-38
                28:  /cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-598-982-38
                29:  /cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-598-982-38
                30:  /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-598-982-38
                31:  /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-598-982-38
                32:  /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-598-982-38
                33:  /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-598-982-38
                34:  /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-598-982-38
                35:  /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-598-982-38
                36:  /cgn2_6/ptodata/2/pna/US098D_COMB.seq:US-09-598-982-38
                37:  /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-598-982-38
                38:  /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-598-982-38
                39:  /cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-598-982-38
                40:  /cgn2_6/ptodata/2/pna/US099D_COMB.seq:US-09-598-982-38
                41:  /cgn2_6/ptodata/2/pna/US099E_COMB.seq:US-09-598-982-38
                42:  /cgn2_6/ptodata/2/pna/US099F_COMB.seq:US-09-598-982-38
                43:  /cgn2_6/ptodata/2/pna/US100A_COMB.seq:US-09-598-982-38
                44:  /cgn2_6/ptodata/2/pna/US100B_COMB.seq:US-09-598-982-38
                45:  /cgn2_6/ptodata/2/pna/US101A_COMB.seq:US-09-598-982-38
                46:  /cgn2_6/ptodata/2/pna/US101B_COMB.seq:US-09-598-982-38
                47:  /cgn2_6/ptodata/2/pna/US102A_COMB.seq:US-09-598-982-38
                48:  /cgn2_6/ptodata/2/pna/US102B_COMB.seq:US-09-598-982-38
                49:  /cgn2_6/ptodata/2/pna/US103A_COMB.seq:US-09-598-982-38
                50:  /cgn2_6/ptodata/2/pna/US103B_COMB.seq:US-09-598-982-38
                51:  /cgn2_6/ptodata/2/pna/US104A_COMB.seq:US-09-598-982-38
                52:  /cgn2_6/ptodata/2/pna/US104B_COMB.seq:US-09-598-982-38
                53:  /cgn2_6/ptodata/2/pna/US106_COMB.seq:US-09-598-982-38
                54:  /cgn2_6/ptodata/2/pna/US107A_COMB.seq:US-09-598-982-38
                55:  /cgn2_6/ptodata/2/pna/US107B_COMB.seq:US-09-598-982-38
                56:  /cgn2_6/ptodata/2/pna/US6000_COMB.seq:US-09-598-982-38
                57:  /cgn2_6/ptodata/2/pna/US6001_COMB.seq:US-09-598-982-38
                58:  /cgn2_6/ptodata/2/pna/US6002_COMB.seq:US-09-598-982-38
                59:  /cgn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-598-982-38
                60:  /cgn2_6/ptodata/2/pna/US6004_COMB.seq:US-09-598-982-38
                61:  /cgn2_6/ptodata/2/pna/US6005_COMB.seq:US-09-598-982-38
                62:  /cgn2_6/ptodata/2/pna/US6006_COMB.seq:US-09-598-982-38
                63:  /cgn2_6/ptodata/2/pna/US6007_COMB.seq:US-09-598-982-38
                64:  /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-09-598-982-38
                65:  /cgn2_6/ptodata/2/pna/US6009_COMB.seq:US-09-598-982-38
                66:  /cgn2_6/ptodata/2/pna/US6010_COMB.seq:US-09-598-982-38
                67:  /cgn2_6/ptodata/2/pna/US6011_COMB.seq:US-09-598-982-38
                68:  /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-09-598-982-38
                69:  /cgn2_6/ptodata/2/pna/US6013_COMB.seq:US-09-598-982-38
                70:  /cgn2_6/ptodata/2/pna/US6014_COMB.seq:US-09-598-982-38
                71:  /cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-09-598-982-38
                72:  /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-09-598-982-38
                73:  /cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-09-598-982-38
                74:  /cgn2_6/ptodata/2/pna/US6018_COMB.seq:US-09-598-982-38
                75:  /cgn2_6/ptodata/2/pna/US6019_COMB.seq:US-09-598-982-38
                76:  /cgn2_6/ptodata/2/pna/US6020_COMB.seq:US-09-598-982-38
                77:  /cgn2_6/ptodata/2/pna/US6021_COMB.seq:US-09-598-982-38
                78:  /cgn2_6/ptodata/2/pna/US6022_COMB.seq:US-09-598-982-38
                79:  /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:US-09-598-982-38
                80:  /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:US-09-598-982-38
                81:  /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-09-598-982-38
                82:  /cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-09-598-982-38
                83:  /cgn2_6/ptodata/2/pna/US6026_COMB.seq:US-09-598-982-38
                84:  /cgn2_6/ptodata/2/pna/US6027_COMB.seq:US-09-598-982-38
                85:  /cgn2_6/ptodata/2/pna/US6028_COMB.seq:US-09-598-982-38
                86:  /cgn2_6/ptodata/2/pna/US6029_COMB.seq:US-09-598-982-38
                87:  /cgn2_6/ptodata/2/pna/US6030_COMB.seq:US-09-598-982-38
                88:  /cgn2_6/ptodata/2/pna/US6031_COMB.seq:US-09-598-982-38
                89:  /cgn2_6/ptodata/2/pna/US6032_COMB.seq:US-09-598-982-38
                90:  /cgn2_6/ptodata/2/pna/US6033_COMB.seq:US-09-598-982-38
                91:  /cgn2_6/ptodata/2/pna/US6034_COMB.seq:US-09-598-982-38
                92:  /cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-09-598-982-38
                93:  /cgn2_6/ptodata/2/pna/US6036_COMB.seq:US-09-598-982-38
                94:  /cgn2_6/ptodata/2/pna/US6037_COMB.seq:US-09-598-982-38
                95:  /cgn2_6/ptodata/2/pna/US6038_COMB.seq:US-09-598-982-38
                96:  /cgn2_6/ptodata/2/pna/US6039_COMB.seq:US-09-598-982-38
                97:  /cgn2_6/ptodata/2/pna/US6040_COMB.seq:US-09-598-982-38
                98:  /cgn2_6/ptodata/2/pna/US6041_COMB.seq:US-09-598-982-38
                99:  /cgn2_6/ptodata/2/pna/US6042_COMB.seq:US-09-598-982-38
                100:  /cgn2_6/ptodata/2/pna/US6043_COMB.seq:US-09-598-982-38
                101:  /cgn2_6/ptodata/2/pna/US6044_COMB.seq:US-09-598-982-38
                102:  /cgn2_6/ptodata/2/pna/US6045_COMB.seq:US-09-598-982-38
                103:  /cgn2_6/ptodata/2/pna/US6046_COMB.seq:US-09-598-982-38
```

```
        104:  /cgn2_6/ptodata/2/pna/US6047_COMB.seq:US-09-598-982-38
        105:  /cgn2_6/ptodata/2/pna/US6048_COMB.seq:US-09-598-982-38
        106:  /cgn2_6/ptodata/2/pna/US6049_COMB.seq:US-09-598-982-38
        107:  /cgn2_6/ptodata/2/pna/US6050_COMB.seq:US-09-598-982-38
        108:  /cgn2_6/ptodata/2/pna/US6051_COMB.seq:US-09-598-982-38
        109:  /cgn2_6/ptodata/2/pna/US6052_COMB.seq:US-09-598-982-38
        110:  /cgn2_6/ptodata/2/pna/US6053_COMB.seq:US-09-598-982-38

  Pred. No. is the number of results predicted by chance to have a
  score greater than or equal to the score of the result being printed,
  and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result             Query
  No.    Score   Match Length DB  ID                      Description
 ----------------------------------------------------------------------------
    1     1396   100.0    771  24  US-09-598-982-38          Sequence 38, Appl

                              ALIGNMENTS

RESULT 1
US-09-598-982-38
; Sequence 38, Application US/09598982
; GENERAL INFORMATION:
;  APPLICANT:  NILES, ANDREW L
;  APPLICANT:  MAFFITT, MARK A
;  APPLICANT:  HAAK-FRENDSCHO, MARY
;  TITLE OF INVENTION:  RECOMBINANT PROTEOLYTIC TYPTASES, ACTIVE SITE MUTANTS
;  TITLE OF INVENTION:  THEREOF, AND METHODS OF MAKING SAME
;  FILE REFERENCE:  CIP TRYPTASE
;  CURRENT APPLICATION NUMBER: US/09/598,982
;  CURRENT FILING DATE:  2000-06-21
;  PRIOR APPLICATION NUMBER:  09/079,970
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  51
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 38
;   LENGTH: 771
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (7)..(753)
US-09-598-982-38

Alignment Scores:
Pred. No.:                0             Length:          771
Score:                    1396.00       Matches:         249
Percent Similarity:       100.00%       Conservative:    0
Best Local Similarity:    100.00%       Mismatches:      0
Query Match:              100.00%       Indels:          0
DB:                       24            Gaps:            0

US-09-598-982-39 (1-249) x US-09-598-982-38 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246
```

```
Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAlaIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValAsnValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

Search completed: July 23, 2004, 08:35:03
Job time : 0.001 secs
```

```
                     GenCore version 5.1.6
             Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 23, 2004, 08:35:57 ; Search time 0.001 Seconds
                                          (without alignments)
                                          383.958 Million cell updates/sec

Title:          US-09-598-982-41
Perfect score:  1398
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1 seqs, 771 residues

Total number of hits satisfying chosen parameters:      2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=Pending_Patents_AA_Main:US-09-598-982-41
-DB=Pending_Patents_NA_Main:US-09-598-982-40 -SUFFIX=pto -OUT=align41_40
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Pending_Patents_NA_Main:US-09-598-982-40
                1:  /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:US-09-598-982-40
                2:  /cgn2_6/ptodata/2/pna/US06_COMB.seq:US-09-598-982-40
                3:  /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-09-598-982-40
                4:  /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-09-598-982-40
                5:  /cgn2_6/ptodata/2/pna/US081_COMB.seq:US-09-598-982-40
                6:  /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-09-598-982-40
                7:  /cgn2_6/ptodata/2/pna/US083_COMB.seq:US-09-598-982-40
                8:  /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-09-598-982-40
                9:  /cgn2_6/ptodata/2/pna/US085_COMB.seq:US-09-598-982-40
                10:  /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-09-598-982-40
                11:  /cgn2_6/ptodata/2/pna/US087_COMB.seq:US-09-598-982-40
                12:  /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-09-598-982-40
                13:  /cgn2_6/ptodata/2/pna/US089_COMB.seq:US-09-598-982-40
                14:  /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-598-982-40
                15:  /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-598-982-40
                16:  /cgn2_6/ptodata/2/pna/US092A_COMB.seq:US-09-598-982-40
                17:  /cgn2_6/ptodata/2/pna/US092B_COMB.seq:US-09-598-982-40
                18:  /cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-598-982-40
                19:  /cgn2_6/ptodata/2/pna/US093B_COMB.seq:US-09-598-982-40
                20:  /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-598-982-40
                21:  /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-598-982-40
                22:  /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-598-982-40
                23:  /cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-598-982-40
                24:  /cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-598-982-40
                25:  /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-598-982-40
                26:  /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-598-982-40
                27:  /cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-598-982-40
                28:  /cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-598-982-40
                29:  /cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-598-982-40
                30:  /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-598-982-40
```

```
31:  /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-598-982-40
32:  /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-598-982-40
33:  /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-598-982-40
34:  /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-598-982-40
35:  /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-598-982-40
36:  /cgn2_6/ptodata/2/pna/US098D_COMB.seq:US-09-598-982-40
37:  /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-598-982-40
38:  /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-598-982-40
39:  /cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-598-982-40
40:  /cgn2_6/ptodata/2/pna/US099D_COMB.seq:US-09-598-982-40
41:  /cgn2_6/ptodata/2/pna/US099E_COMB.seq:US-09-598-982-40
42:  /cgn2_6/ptodata/2/pna/US099F_COMB.seq:US-09-598-982-40
43:  /cgn2_6/ptodata/2/pna/US100A_COMB.seq:US-09-598-982-40
44:  /cgn2_6/ptodata/2/pna/US100B_COMB.seq:US-09-598-982-40
45:  /cgn2_6/ptodata/2/pna/US101A_COMB.seq:US-09-598-982-40
46:  /cgn2_6/ptodata/2/pna/US101B_COMB.seq:US-09-598-982-40
47:  /cgn2_6/ptodata/2/pna/US102A_COMB.seq:US-09-598-982-40
48:  /cgn2_6/ptodata/2/pna/US102B_COMB.seq:US-09-598-982-40
49:  /cgn2_6/ptodata/2/pna/US103A_COMB.seq:US-09-598-982-40
50:  /cgn2_6/ptodata/2/pna/US103B_COMB.seq:US-09-598-982-40
51:  /cgn2_6/ptodata/2/pna/US104A_COMB.seq:US-09-598-982-40
52:  /cgn2_6/ptodata/2/pna/US104B_COMB.seq:US-09-598-982-40
53:  /cgn2_6/ptodata/2/pna/US106_COMB.seq:US-09-598-982-40
54:  /cgn2_6/ptodata/2/pna/US107A_COMB.seq:US-09-598-982-40
55:  /cgn2_6/ptodata/2/pna/US107B_COMB.seq:US-09-598-982-40
56:  /cgn2_6/ptodata/2/pna/US6000_COMB.seq:US-09-598-982-40
57:  /cgn2_6/ptodata/2/pna/US6001_COMB.seq:US-09-598-982-40
58:  /cgn2_6/ptodata/2/pna/US6002_COMB.seq:US-09-598-982-40
59:  /cgn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-598-982-40
60:  /cgn2_6/ptodata/2/pna/US6004_COMB.seq:US-09-598-982-40
61:  /cgn2_6/ptodata/2/pna/US6005_COMB.seq:US-09-598-982-40
62:  /cgn2_6/ptodata/2/pna/US6006_COMB.seq:US-09-598-982-40
63:  /cgn2_6/ptodata/2/pna/US6007_COMB.seq:US-09-598-982-40
64:  /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-09-598-982-40
65:  /cgn2_6/ptodata/2/pna/US6009_COMB.seq:US-09-598-982-40
66:  /cgn2_6/ptodata/2/pna/US6010_COMB.seq:US-09-598-982-40
67:  /cgn2_6/ptodata/2/pna/US6011_COMB.seq:US-09-598-982-40
68:  /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-09-598-982-40
69:  /cgn2_6/ptodata/2/pna/US6013_COMB.seq:US-09-598-982-40
70:  /cgn2_6/ptodata/2/pna/US6014_COMB.seq:US-09-598-982-40
71:  /cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-09-598-982-40
72:  /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-09-598-982-40
73:  /cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-09-598-982-40
74:  /cgn2_6/ptodata/2/pna/US6018_COMB.seq:US-09-598-982-40
75:  /cgn2_6/ptodata/2/pna/US6019_COMB.seq:US-09-598-982-40
76:  /cgn2_6/ptodata/2/pna/US6020_COMB.seq:US-09-598-982-40
77:  /cgn2_6/ptodata/2/pna/US6021_COMB.seq:US-09-598-982-40
78:  /cgn2_6/ptodata/2/pna/US6022_COMB.seq:US-09-598-982-40
79:  /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:US-09-598-982-40
80:  /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:US-09-598-982-40
81:  /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-09-598-982-40
82:  /cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-09-598-982-40
83:  /cgn2_6/ptodata/2/pna/US6026_COMB.seq:US-09-598-982-40
84:  /cgn2_6/ptodata/2/pna/US6027_COMB.seq:US-09-598-982-40
85:  /cgn2_6/ptodata/2/pna/US6028_COMB.seq:US-09-598-982-40
86:  /cgn2_6/ptodata/2/pna/US6029_COMB.seq:US-09-598-982-40
87:  /cgn2_6/ptodata/2/pna/US6030_COMB.seq:US-09-598-982-40
88:  /cgn2_6/ptodata/2/pna/US6031_COMB.seq:US-09-598-982-40
89:  /cgn2_6/ptodata/2/pna/US6032_COMB.seq:US-09-598-982-40
90:  /cgn2_6/ptodata/2/pna/US6033_COMB.seq:US-09-598-982-40
91:  /cgn2_6/ptodata/2/pna/US6034_COMB.seq:US-09-598-982-40
92:  /cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-09-598-982-40
93:  /cgn2_6/ptodata/2/pna/US6036_COMB.seq:US-09-598-982-40
94:  /cgn2_6/ptodata/2/pna/US6037_COMB.seq:US-09-598-982-40
95:  /cgn2_6/ptodata/2/pna/US6038_COMB.seq:US-09-598-982-40
96:  /cgn2_6/ptodata/2/pna/US6039_COMB.seq:US-09-598-982-40
97:  /cgn2_6/ptodata/2/pna/US6040_COMB.seq:US-09-598-982-40
98:  /cgn2_6/ptodata/2/pna/US6041_COMB.seq:US-09-598-982-40
99:  /cgn2_6/ptodata/2/pna/US6042_COMB.seq:US-09-598-982-40
100:  /cgn2_6/ptodata/2/pna/US6043_COMB.seq:US-09-598-982-40
101:  /cgn2_6/ptodata/2/pna/US6044_COMB.seq:US-09-598-982-40
102:  /cgn2_6/ptodata/2/pna/US6045_COMB.seq:US-09-598-982-40
103:  /cgn2_6/ptodata/2/pna/US6046_COMB.seq:US-09-598-982-40
```

```
          104:  /cgn2_6/ptodata/2/pna/US6047_COMB.seq:US-09-598-982-40
          105:  /cgn2_6/ptodata/2/pna/US6048_COMB.seq:US-09-598-982-40
          106:  /cgn2_6/ptodata/2/pna/US6049_COMB.seq:US-09-598-982-40
          107:  /cgn2_6/ptodata/2/pna/US6050_COMB.seq:US-09-598-982-40
          108:  /cgn2_6/ptodata/2/pna/US6051_COMB.seq:US-09-598-982-40
          109:  /cgn2_6/ptodata/2/pna/US6052_COMB.seq:US-09-598-982-40
          110:  /cgn2_6/ptodata/2/pna/US6053_COMB.seq:US-09-598-982-40

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result            Query
  No.   Score     Match Length DB  ID                         Description
 --------------------------------------------------------------------------
    1    1398    100.0    771  24  US-09-598-982-40          Sequence 40, Appl

                              ALIGNMENTS

RESULT 1
US-09-598-982-40
; Sequence 40, Application US/09598982
; GENERAL INFORMATION:
;  APPLICANT:  NILES, ANDREW L
;  APPLICANT:  MAFFITT, MARK A
;  APPLICANT:  HAAK-FRENDSCHO, MARY
;  TITLE OF INVENTION:  RECOMBINANT PROTEOLYTIC TYPTASES, ACTIVE SITE MUTANTS
;  TITLE OF INVENTION:  THEREOF, AND METHODS OF MAKING SAME
;  FILE REFERENCE:  CIP TRYPTASE
;  CURRENT APPLICATION NUMBER: US/09/598,982
;  CURRENT FILING DATE:  2000-06-21
;  PRIOR APPLICATION NUMBER:  09/079,970
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  51
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 40
;   LENGTH: 771
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (7)..(753)
US-09-598-982-40

Alignment Scores:
Pred. No.:              0               Length:         771
Score:                  1398.00         Matches:        249
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     24              Gaps:           0

US-09-598-982-41 (1-249) x US-09-598-982-40 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValAsnValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspAlaGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCCGGCGGA 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCTCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

Search completed: July 23, 2004, 08:35:57
Job time : 0.001 secs
```

```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 23, 2004, 08:36:46 ; Search time 0.001 Seconds
                                           (without alignments)
                                           383.958 Million cell updates/sec

Title:          US-09-598-982-43
Perfect score:  1398
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       1 seqs, 771 residues

Total number of hits satisfying chosen parameters:      2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 1 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=soft -Q=Pending_Patents_AA_Main:US-09-598-982-43
-DB=Pending_Patents_NA_Main:US-09-598-982-42 -SUFFIX=pto -OUT=align43_42
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=1 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=1 -MODE=LOCAL -OUTFMT=pto -NORM=ext
-HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -NCPU=6 -NO_XLPXY -NEG_SCORES=0
-LONGLOG -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7 -YGAPOP=10
-YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Pending_Patents_NA_Main:US-09-598-982-42
                1:  /cgn2_6/ptodata/2/pna/PCTUS_COMB.seq:US-09-598-982-42
                2:  /cgn2_6/ptodata/2/pna/US06_COMB.seq:US-09-598-982-42
                3:  /cgn2_6/ptodata/2/pna/US07_COMB.seq:US-09-598-982-42
                4:  /cgn2_6/ptodata/2/pna/US080_COMB.seq:US-09-598-982-42
                5:  /cgn2_6/ptodata/2/pna/US081_COMB.seq:US-09-598-982-42
                6:  /cgn2_6/ptodata/2/pna/US082_COMB.seq:US-09-598-982-42
                7:  /cgn2_6/ptodata/2/pna/US083_COMB.seq:US-09-598-982-42
                8:  /cgn2_6/ptodata/2/pna/US084_COMB.seq:US-09-598-982-42
                9:  /cgn2_6/ptodata/2/pna/US085_COMB.seq:US-09-598-982-42
                10:  /cgn2_6/ptodata/2/pna/US086_COMB.seq:US-09-598-982-42
                11:  /cgn2_6/ptodata/2/pna/US087_COMB.seq:US-09-598-982-42
                12:  /cgn2_6/ptodata/2/pna/US088_COMB.seq:US-09-598-982-42
                13:  /cgn2_6/ptodata/2/pna/US089_COMB.seq:US-09-598-982-42
                14:  /cgn2_6/ptodata/2/pna/US090_COMB.seq:US-09-598-982-42
                15:  /cgn2_6/ptodata/2/pna/US091_COMB.seq:US-09-598-982-42
                16:  /cgn2_6/ptodata/2/pna/US092A_COMB.seq:US-09-598-982-42
                17:  /cgn2_6/ptodata/2/pna/US092B_COMB.seq:US-09-598-982-42
                18:  /cgn2_6/ptodata/2/pna/US093A_COMB.seq:US-09-598-982-42
                19:  /cgn2_6/ptodata/2/pna/US093B_COMB.seq:US-09-598-982-42
                20:  /cgn2_6/ptodata/2/pna/US094_COMB.seq:US-09-598-982-42
                21:  /cgn2_6/ptodata/2/pna/US095A_COMB.seq:US-09-598-982-42
                22:  /cgn2_6/ptodata/2/pna/US095B_COMB.seq:US-09-598-982-42
                23:  /cgn2_6/ptodata/2/pna/US095C_COMB.seq:US-09-598-982-42
                24:  /cgn2_6/ptodata/2/pna/US095D_COMB.seq:US-09-598-982-42
                25:  /cgn2_6/ptodata/2/pna/US096A_COMB.seq:US-09-598-982-42
                26:  /cgn2_6/ptodata/2/pna/US096B_COMB.seq:US-09-598-982-42
                27:  /cgn2_6/ptodata/2/pna/US096C_COMB.seq:US-09-598-982-42
                28:  /cgn2_6/ptodata/2/pna/US096D_COMB.seq:US-09-598-982-42
                29:  /cgn2_6/ptodata/2/pna/US096E_COMB.seq:US-09-598-982-42
                30:  /cgn2_6/ptodata/2/pna/US097A_COMB.seq:US-09-598-982-42
                31:  /cgn2_6/ptodata/2/pna/US097B_COMB.seq:US-09-598-982-42
                32:  /cgn2_6/ptodata/2/pna/US097C_COMB.seq:US-09-598-982-42
                33:  /cgn2_6/ptodata/2/pna/US098A_COMB.seq:US-09-598-982-42
                34:  /cgn2_6/ptodata/2/pna/US098B_COMB.seq:US-09-598-982-42
                35:  /cgn2_6/ptodata/2/pna/US098C_COMB.seq:US-09-598-982-42
                36:  /cgn2_6/ptodata/2/pna/US098D_COMB.seq:US-09-598-982-42
                37:  /cgn2_6/ptodata/2/pna/US099A_COMB.seq:US-09-598-982-42
                38:  /cgn2_6/ptodata/2/pna/US099B_COMB.seq:US-09-598-982-42
                39:  /cgn2_6/ptodata/2/pna/US099C_COMB.seq:US-09-598-982-42
                40:  /cgn2_6/ptodata/2/pna/US099D_COMB.seq:US-09-598-982-42
                41:  /cgn2_6/ptodata/2/pna/US099E_COMB.seq:US-09-598-982-42
                42:  /cgn2_6/ptodata/2/pna/US099F_COMB.seq:US-09-598-982-42
                43:  /cgn2_6/ptodata/2/pna/US100A_COMB.seq:US-09-598-982-42
                44:  /cgn2_6/ptodata/2/pna/US100B_COMB.seq:US-09-598-982-42
                45:  /cgn2_6/ptodata/2/pna/US101A_COMB.seq:US-09-598-982-42
                46:  /cgn2_6/ptodata/2/pna/US101B_COMB.seq:US-09-598-982-42
                47:  /cgn2_6/ptodata/2/pna/US102A_COMB.seq:US-09-598-982-42
                48:  /cgn2_6/ptodata/2/pna/US102B_COMB.seq:US-09-598-982-42
                49:  /cgn2_6/ptodata/2/pna/US103A_COMB.seq:US-09-598-982-42
                50:  /cgn2_6/ptodata/2/pna/US103B_COMB.seq:US-09-598-982-42
                51:  /cgn2_6/ptodata/2/pna/US104A_COMB.seq:US-09-598-982-42
                52:  /cgn2_6/ptodata/2/pna/US104B_COMB.seq:US-09-598-982-42
                53:  /cgn2_6/ptodata/2/pna/US106_COMB.seq:US-09-598-982-42
                54:  /cgn2_6/ptodata/2/pna/US107A_COMB.seq:US-09-598-982-42
                55:  /cgn2_6/ptodata/2/pna/US107B_COMB.seq:US-09-598-982-42
                56:  /cgn2_6/ptodata/2/pna/US6000_COMB.seq:US-09-598-982-42
                57:  /cgn2_6/ptodata/2/pna/US6001_COMB.seq:US-09-598-982-42
                58:  /cgn2_6/ptodata/2/pna/US6002_COMB.seq:US-09-598-982-42
                59:  /cgn2_6/ptodata/2/pna/US6003_COMB.seq:US-09-598-982-42
                60:  /cgn2_6/ptodata/2/pna/US6004_COMB.seq:US-09-598-982-42
                61:  /cgn2_6/ptodata/2/pna/US6005_COMB.seq:US-09-598-982-42
                62:  /cgn2_6/ptodata/2/pna/US6006_COMB.seq:US-09-598-982-42
                63:  /cgn2_6/ptodata/2/pna/US6007_COMB.seq:US-09-598-982-42
                64:  /cgn2_6/ptodata/2/pna/US6008_COMB.seq:US-09-598-982-42
                65:  /cgn2_6/ptodata/2/pna/US6009_COMB.seq:US-09-598-982-42
                66:  /cgn2_6/ptodata/2/pna/US6010_COMB.seq:US-09-598-982-42
                67:  /cgn2_6/ptodata/2/pna/US6011_COMB.seq:US-09-598-982-42
                68:  /cgn2_6/ptodata/2/pna/US6012_COMB.seq:US-09-598-982-42
                69:  /cgn2_6/ptodata/2/pna/US6013_COMB.seq:US-09-598-982-42
                70:  /cgn2_6/ptodata/2/pna/US6014_COMB.seq:US-09-598-982-42
                71:  /cgn2_6/ptodata/2/pna/US6015_COMB.seq:US-09-598-982-42
                72:  /cgn2_6/ptodata/2/pna/US6016_COMB.seq:US-09-598-982-42
                73:  /cgn2_6/ptodata/2/pna/US6017_COMB.seq:US-09-598-982-42
                74:  /cgn2_6/ptodata/2/pna/US6018_COMB.seq:US-09-598-982-42
                75:  /cgn2_6/ptodata/2/pna/US6019_COMB.seq:US-09-598-982-42
                76:  /cgn2_6/ptodata/2/pna/US6020_COMB.seq:US-09-598-982-42
                77:  /cgn2_6/ptodata/2/pna/US6021_COMB.seq:US-09-598-982-42
                78:  /cgn2_6/ptodata/2/pna/US6022_COMB.seq:US-09-598-982-42
                79:  /cgn2_6/ptodata/2/pna/US6023A_COMB.seq:US-09-598-982-42
                80:  /cgn2_6/ptodata/2/pna/US6023B_COMB.seq:US-09-598-982-42
                81:  /cgn2_6/ptodata/2/pna/US6024_COMB.seq:US-09-598-982-42
                82:  /cgn2_6/ptodata/2/pna/US6025_COMB.seq:US-09-598-982-42
                83:  /cgn2_6/ptodata/2/pna/US6026_COMB.seq:US-09-598-982-42
                84:  /cgn2_6/ptodata/2/pna/US6027_COMB.seq:US-09-598-982-42
                85:  /cgn2_6/ptodata/2/pna/US6028_COMB.seq:US-09-598-982-42
                86:  /cgn2_6/ptodata/2/pna/US6029_COMB.seq:US-09-598-982-42
                87:  /cgn2_6/ptodata/2/pna/US6030_COMB.seq:US-09-598-982-42
                88:  /cgn2_6/ptodata/2/pna/US6031_COMB.seq:US-09-598-982-42
                89:  /cgn2_6/ptodata/2/pna/US6032_COMB.seq:US-09-598-982-42
                90:  /cgn2_6/ptodata/2/pna/US6033_COMB.seq:US-09-598-982-42
                91:  /cgn2_6/ptodata/2/pna/US6034_COMB.seq:US-09-598-982-42
                92:  /cgn2_6/ptodata/2/pna/US6035_COMB.seq:US-09-598-982-42
                93:  /cgn2_6/ptodata/2/pna/US6036_COMB.seq:US-09-598-982-42
                94:  /cgn2_6/ptodata/2/pna/US6037_COMB.seq:US-09-598-982-42
                95:  /cgn2_6/ptodata/2/pna/US6038_COMB.seq:US-09-598-982-42
                96:  /cgn2_6/ptodata/2/pna/US6039_COMB.seq:US-09-598-982-42
                97:  /cgn2_6/ptodata/2/pna/US6040_COMB.seq:US-09-598-982-42
                98:  /cgn2_6/ptodata/2/pna/US6041_COMB.seq:US-09-598-982-42
                99:  /cgn2_6/ptodata/2/pna/US6042_COMB.seq:US-09-598-982-42
                100:  /cgn2_6/ptodata/2/pna/US6043_COMB.seq:US-09-598-982-42
                101:  /cgn2_6/ptodata/2/pna/US6044_COMB.seq:US-09-598-982-42
                102:  /cgn2_6/ptodata/2/pna/US6045_COMB.seq:US-09-598-982-42
                103:  /cgn2_6/ptodata/2/pna/US6046_COMB.seq:US-09-598-982-42
```

```
          104:  /cgn2_6/ptodata/2/pna/US6047_COMB.seq:US-09-598-982-42
          105:  /cgn2_6/ptodata/2/pna/US6048_COMB.seq:US-09-598-982-42
          106:  /cgn2_6/ptodata/2/pna/US6049_COMB.seq:US-09-598-982-42
          107:  /cgn2_6/ptodata/2/pna/US6050_COMB.seq:US-09-598-982-42
          108:  /cgn2_6/ptodata/2/pna/US6051_COMB.seq:US-09-598-982-42
          109:  /cgn2_6/ptodata/2/pna/US6052_COMB.seq:US-09-598-982-42
          110:  /cgn2_6/ptodata/2/pna/US6053_COMB.seq:US-09-598-982-42

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                     SUMMARIES
                       %
Result                Query
  No.     Score    Match Length DB  ID                              Description
 ------------------------------------------------------------------------------
    1     1398    100.0    771  24  US-09-598-982-42                  Sequence 42, Appl

                                     ALIGNMENTS

RESULT 1
US-09-598-982-42
; Sequence 42, Application US/09598982
; GENERAL INFORMATION:
;  APPLICANT:  NILES, ANDREW L
;  APPLICANT:  MAFFITT, MARK A
;  APPLICANT:  HAAK-FRENDSCHO, MARY
;  TITLE OF INVENTION:  RECOMBINANT PROTEOLYTIC TYPTASES, ACTIVE SITE MUTANTS
;  TITLE OF INVENTION:  THEREOF, AND METHODS OF MAKING SAME
;  FILE REFERENCE:  CIP TRYPTASE
;  CURRENT APPLICATION NUMBER: US/09/598,982
;  CURRENT FILING DATE:  2000-06-21
;  PRIOR APPLICATION NUMBER:  09/079,970
;  PRIOR FILING DATE:  1998-04-15
;  NUMBER OF SEQ ID NOS:  51
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 42
;   LENGTH: 771
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (7)..(753)
US-09-598-982-42

Alignment Scores:
Pred. No.:                0              Length:        771
Score:                    1398.00        Matches:       249
Percent Similarity:       100.00%        Conservative:  0
Best Local Similarity:    100.00%        Mismatches:    0
Query Match:              100.00%        Indels:        0
DB:                       24             Gaps:          0

US-09-598-982-43 (1-249) x US-09-598-982-42 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaHisCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValAsnValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspAlaGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCCGGCGGA 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCACTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

Search completed: July 23, 2004, 08:36:47
Job time : 1 secs
```

```
                    GenCore version 5.1.6
             Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          July 22, 2004, 13:26:16 ; Search time 52 Seconds
                                           (without alignments)
                                           249.336 Million cell updates/sec

Title:           US-09-598-982-21
Perfect score:   1393
Sequence:        1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters:       141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SwissProt_42:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                    %
Result            Query
  No.    Score    Match  Length  DB  ID                      Description
---------------------------------------------------------------------------
   1      1368     98.2     275   1  TRB2_HUMAN              P20231 homo sapien
   2      1363     97.8     275   1  TRB1_HUMAN              Q15661 homo sapien
   3      1282     92.0     275   1  TRYA_HUMAN              P15157 homo sapien
   4      1091     78.3     276   1  MCT6_MOUSE              P21845 mus musculu
   5      1077     77.3     275   1  TRYT_CANFA              P15944 canis famil
   6      1075     77.2     270   1  TRYT_MERUN              P50342 meriones un
   7      1054     75.7     273   1  MCT7_MOUSE              Q02844 mus musculu
   8      1051     75.4     274   1  MCT6_RAT                P50343 rattus norv
   9      1037     74.4     273   1  MCT7_RAT                P27435 rattus norv
  10      1030     73.9     273   1  TRYT_SHEEP              Q9xsm2 ovis aries
  11      1006     72.2     275   1  TRYT_PIG                Q9n2d1 sus scrofa
  12       988     70.9     235   1  TRYD_HUMAN              Q9bzj3 homo sapien
  13     684.5     49.1     269   1  TRYM_CANFA              P19236 canis famil
  14     605.5     43.5     311   1  TRYG_MOUSE              Q9qul7 mus musculu
  15     593.5     42.6     321   1  TRYG_HUMAN              Q9nrr2 homo sapien
  16     563.5     40.5     290   1  PR27_HUMAN              Q9bqr3 homo sapien
  17       529     38.0     342   1  PSS8_RAT                Q9es87 rattus norv
  18       526     37.8     342   1  PSS8_MOUSE              Q9esd1 mus musculu
  19       512     36.8     343   1  PSS8_HUMAN              Q16651 homo sapien
  20     501.5     36.0     317   1  BSS4_HUMAN              Q9gzn4 homo sapien
  21     493.5     35.4     314   1  TEST_HUMAN              Q9y6m0 homo sapien
  22       491     35.2     453   1  TMS3_MOUSE              Q8k1t0 mus musculu
  23       488     35.0     455   1  TMS5_MOUSE              Q9er04 mus musculu
  24       476     34.2     324   1  TEST_MOUSE              Q9jhj7 mus musculu
  25     474.5     34.1     454   1  TMS3_HUMAN              P57727 homo sapien
  26       467     33.5     625   1  FA11_HUMAN              P03951 homo sapien
  27     464.5     33.3     811   1  TMS6_MOUSE              Q9dbi0 mus musculu
  28       463     33.2     306   1  BSS4_MOUSE              Q9er10 mus musculu
  29     460.5     33.1     418   1  HATT_HUMAN              O60235 homo sapien
  30     460.5     33.1     811   1  TMS6_HUMAN              Q8iu80 homo sapien
  31       457     32.8     457   1  TMS5_HUMAN              Q9h3s3 homo sapien
  32       451     32.4     638   1  KAL_HUMAN               P03952 homo sapien
  33       448     32.2     338   1  PLMN_HORSE              P80010 equus cabal
  34       447     32.1     638   1  KAL_MOUSE               P26262 mus musculu
  35     446.5     32.1     812   1  PLMN_BOVIN              P06868 bos taurus
  36       445     31.9     638   1  KAL_RAT                 P14272 rattus norv
  37     444.5     31.9     271   1  EL2_RAT                 P00774 rattus norv
  38       444     31.9     333   1  PLMN_CANFA              P80009 canis famil
  39       443     31.8     245   1  CTRB_BOVIN              P00767 bos taurus
  40       441     31.7     343   1  PLMN_SHEEP              P81286 ovis aries
  41     436.5     31.3     269   1  EL2A_HUMAN              P08217 homo sapien
  42       434     31.2     490   1  TMS2_MOUSE              Q9jiq8 mus musculu
  43     430.5     30.9     271   1  EL2_MOUSE               P05208 mus musculu
  44       429     30.8     435   1  TMS4_MOUSE              Q8vca5 mus musculu
  45       429     30.8     810   1  PLMN_HUMAN              P00747 homo sapien
```

```
                              ALIGNMENTS

RESULT 1
TRB2_HUMAN
ID   TRB2_HUMAN     STANDARD;      PRT;   275 AA.
AC   P20231; O95827; Q15664; Q9UQI6; Q9UQI7;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tryptase beta-2 precursor (EC 3.4.21.59) (Tryptase 2) (Tryptase II).
GN   TPSB2 OR TPS2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (VARIANT BETA-2).
RC   TISSUE=Lung;
RX   MEDLINE=90369005; PubMed=2203827;
RA   Miller J.S., Moxley G., Schwartz L.B.;
RT   "Cloning and characterization of a second complementary DNA for human
RT   tryptase.";
RL   J. Clin. Invest. 86:864-870(1990).
RN   [2]
RP   SEQUENCE FROM N.A. (VARIANT BETA-2).
RX   MEDLINE=93166209; PubMed=8434231;
RA   Blom T., Hellman L.;
RT   "Characterization of a tryptase mRNA expressed in the human basophil
RT   cell line KU812.";
RL   Scand. J. Immunol. 37:203-208(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (VARIANTS BETA-2 AND BETA-3).
RX   MEDLINE=90251647; PubMed=2187193;
RA   Vanderslice P., Ballinger S.M., Tam E.K., Goldstein S.M., Craik C.S.,
RA   Caughey G.H.;
RT   "Human mast cell tryptase: multiple cDNAs and genes reveal a
RT   multigene serine protease family.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:3811-3815(1990).
RN   [4]
RP   SEQUENCE FROM N.A. (VARIANTS BETA-2 AND BETA-3).
RX   MEDLINE=99121069; PubMed=9920877;
RA   Pallaoro M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
RT   "Characterization of genes encoding known and novel human mast cell
RT   tryptases on chromosome 16p13.3.";
RL   J. Biol. Chem. 274:3355-3362(1999).
RN   [5]
RP   REVISIONS.
RA   Pallaoro M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
```

```
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RX   MEDLINE=98180625; PubMed=9521329;
RA   Pereira P.J.B., Bergner A., Macedo-Ribeiro S., Huber R.,
RA   Matschiner G., Fritz H., Sommerhoff C.P., Bode W.;
RT   "Human beta-tryptase is a ring-like tetramer with active sites facing
RT   a central pore.";
RL   Nature 392:306-311(1998).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS).
RX   MEDLINE=99432168; PubMed=10500112;
RA   Sommerhoff C.P., Bode W., Pereira P.J.B., Stubbs M.T.,
RA   Stuerzebecher J., Piechottka G.P., Matschiner G., Bergner A.;
RT   "The structure of the human betaII-tryptase tetramer: fo(u)r better or
RT   worse.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:10984-10991(1999).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
CC   -!- POLYMORPHISM: There are two alleles; beta-II and beta-III which
CC       differ by 3 residues.
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M37488; AAA51843.1; -.
DR   EMBL; M33492; AAA36779.1; -.
DR   EMBL; M33493; AAA36780.1; -.
DR   EMBL; S55551; AAD13876.1; -.
DR   EMBL; AF099143; AAD17859.2; -.
DR   EMBL; AF099145; AAD17857.1; -.
DR   EMBL; AF099146; AAD17858.1; -.
DR   EMBL; BC029356; AAH29356.1; -.
DR   PIR; B35863; B35863.
DR   PDB; 1AAO; 31-JAN-94.
DR   PDB; 1A0L; 23-MAR-99.
DR   MEROPS; S01.027; -.
DR   MEROPS; S01.242; -.
DR   Genew; HGNC:14120; TPSB2.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
```

```
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;
KW   Polymorphism; 3D-structure.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     30       ACTIVATION PEPTIDE.
FT   CHAIN        31    275       TRYPTASE BETA-2.
FT   ACT_SITE     74     74       CHARGE RELAY SYSTEM.
FT   ACT_SITE    121    121       CHARGE RELAY SYSTEM.
FT   ACT_SITE    224    224       CHARGE RELAY SYSTEM.
FT   DISULFID     59     75
FT   DISULFID    155    230
FT   DISULFID    188    211
FT   DISULFID    220    248
FT   CARBOHYD    233    233       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      51     53       HGP -> RDR (IN BETA-III).
FT                                /FTId=VAR_012104.
FT   CONFLICT    132    132       K -> N (IN REF. 3).
SQ   SEQUENCE   275 AA;  30529 MW;  2B27396C51F5C7A0 CRC64;

  Query Match             98.2%;  Score 1368;  DB 1;  Length 275;
  Best Local Similarity   99.6%;  Pred. No. 2.6e-118;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275


RESULT 2
TRB1_HUMAN
ID   TRB1_HUMAN     STANDARD;      PRT;   275 AA.
AC   Q15661; Q15663; Q9H2Y4;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tryptase beta-1 precursor (EC 3.4.21.59) (Tryptase 1) (Tryptase I).
GN   TPSB1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90251647; PubMed=2187193;
RA   Vanderslice P., Ballinger S.M., Tam E.K., Goldstein S.M., Craik C.S.,
RA   Caughey G.H.;
RT   "Human mast cell tryptase: multiple cDNAs and genes reveal a multigene
RT   serine protease family.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:3811-3815(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99121069; PubMed=9920877;
RA   Pallaoro M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
RT   "Characterization of genes encoding known and novel human mast cell
RT   tryptases on chromosome 16p13.3.";
```

```
RL   J. Biol. Chem. 274:3355-3362(1999).
RN   [3]
RP   SEQUENCE OF 54-275 FROM N.A. (ISOFORM 2).
RA   Wang H.W., McNeil H.P., Thomas P.S., Murphy B.N., Webster M.J.,
RA   Hettiaratchi A., King G., Heywood G.J., Huang C., Stevens R.L.,
RA   Hunt J.E.;
RT   "Molecular cloning and characterization of novel human tryptase cDNAs
RT   and splicing variants.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q15661-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q15661-2; Sequence=VSP_005375;
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M33494; AAC83172.1; -.
DR   EMBL; M33491; AAA36778.1; -.
DR   EMBL; AF099144; AAD17860.1; -.
DR   EMBL; AF206667; AAG35697.1; -.
DR   PIR; A35863; A35863.
DR   HSSP; P20231; 1A0L.
DR   MEROPS; S01.242; -.
DR   Genew; HGNC:12019; TPSB1.
DR   MIM; 191081; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;
KW   Alternative splicing; Polymorphism.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     30       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   CHAIN        31    275       TRYPTASE BETA-1.
FT   ACT_SITE     74     74       CHARGE RELAY SYSTEM.
FT   ACT_SITE    121    121       CHARGE RELAY SYSTEM.
FT   ACT_SITE    224    224       CHARGE RELAY SYSTEM.
FT   DISULFID     59     75       BY SIMILARITY.
FT   DISULFID    155    230       BY SIMILARITY.
FT   DISULFID    188    211       BY SIMILARITY.
FT   DISULFID    220    248       BY SIMILARITY.
FT   CARBOHYD    132    132       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    233    233       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     79     87       Missing (in isoform 2).
FT                                /FTId=VSP_005375.
FT   VARIANT      18     18       A -> V (in dbSNP:1800984).
FT                                /FTId=VAR_014557.
FT   VARIANT      23     23       G -> V (in dbSNP:1800986).
FT                                /FTId=VAR_014558.
FT   VARIANT      85     85       A -> T (in dbSNP:2234902).
FT                                /FTId=VAR_014559.
FT   VARIANT     132    132       N -> K (in dbSNP:1800991).
FT                                /FTId=VAR_016102.
FT   VARIANT     141    141       T -> A (in dbSNP:1800992).
FT                                /FTId=VAR_014560.
FT   VARIANT     162    162       D -> N (in dbSNP:2234641).
FT                                /FTId=VAR_014561.
FT   VARIANT     170    170       P -> S (in dbSNP:2234904).
FT                                /FTId=VAR_014562.
FT   VARIANT     215    215       T -> S (in dbSNP:2234905).
FT                                /FTId=VAR_014563.
FT   VARIANT     216    216       R -> Q (in dbSNP:2234906).
FT                                /FTId=VAR_014564.
SQ   SEQUENCE   275 AA;  30515 MW;  ADC48FDC51F37112 CRC64;

  Query Match             97.8%;  Score 1363;  DB 1;  Length 275;
  Best Local Similarity   99.2%;  Pred. No. 7.4e-118;
  Matches  243;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275


RESULT 3
TRYA_HUMAN
ID   TRYA_HUMAN     STANDARD;      PRT;   275 AA.
AC   P15157; Q9H2Y5; Q9UQI1;
DT   01-APR-1990 (Rel. 14, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Alpha-tryptase precursor (EC 3.4.21.59) (Tryptase 1).
GN   TPS1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RC   TISSUE=Lung;
RX   MEDLINE=90009311; PubMed=2677049;
RA   Miller J.S., Westin E.H., Schwartz L.B.;
RT   "Cloning and characterization of complementary DNA for human
RT   tryptase.";
RL   J. Clin. Invest. 84:1188-1195(1989).
RN   [2]
RP   REVISIONS TO 89-93 AND 108.
RA   Schwartz L.B.;
RL   Submitted (MAR-1990) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM 1).
RX   MEDLINE=99121069; PubMed=9920877;
RA   Pallaoro M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
RT   "Characterization of genes encoding known and novel human mast cell
RT   tryptases on chromosome 16p13.3.";
RL   J. Biol. Chem. 274:3355-3362(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC   TISSUE=Lung;
```

```
RA   Wang H.W., McNeil H.P., Thomas P.S., Murphy B.N., Webster M.J.,
RA   Hettiaratchi A., King G., Heywood G.J., Huang C., Stevens R.L.,
RA   Hunt J.E.;
RT   "Molecular cloning and characterization of novel human tryptase cDNAs
RT   and splicing variants.";
RL   Submitted (NOV-1999) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 31-50, AND PITUITARY, SEQUENCE OF 31-38.
RC   TISSUE=Lung;
RX   MEDLINE=87109258; PubMed=3543004;
RA   Cromlish J.A., Seidah N.G., Marcinkiewcz M., Hamelin J., Johnson D.A.,
RA   Chretein M.;
RT   "Human pituitary tryptase: molecular forms, NH2-terminal sequence,
RT   immunocytochemical localization, and specificity with prohormone and
RT   fluorogenic substrates.";
RL   J. Biol. Chem. 262:1363-1373(1987).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=P15157-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=P15157-2; Sequence=VSP_005374;
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M30038; AAA86934.1; -.
DR   EMBL; AF098328; AAD17846.1; -.
DR   EMBL; AF206665; AAG35695.1; -.
DR   EMBL; AF206666; AAG35696.1; -.
DR   HSSP; P20231; 1A0L.
DR   MEROPS; S01.015; -.
DR   MEROPS; S01.143; -.
DR   Genew; HGNC:12018; TPS1.
DR   MIM; 191080; -.
DR   GO; GO:0008236; F:serine-type peptidase activity; TAS.
DR   GO; GO:0006952; P:defense response; TAS.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;
KW   Polymorphism; Alternative splicing.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     30       ACTIVATION PEPTIDE.
FT   CHAIN        31    275       ALPHA-TRYPTASE.
FT   ACT_SITE     74     74       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    121    121       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    224    224       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     59     75       BY SIMILARITY.
FT   DISULFID    155    230       BY SIMILARITY.
FT   DISULFID    188    211       BY SIMILARITY.
FT   DISULFID    220    248       BY SIMILARITY.
FT   CARBOHYD    132    132       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    233    233       N-LINKED (GLCNAC. . .) (POTENTIAL).
```

```
FT   VARSPLIC     79     87       Missing (in isoform 2).
FT                                /FTId=VSP_005374.
FT   VARIANT      15     15       R -> P (IN ALPHA-II).
FT                                /FTId=VAR_012102.
FT   VARIANT     221    221       K -> Q (IN ALPHA-II; dbSNP:1137382).
FT                                /FTId=VAR_012103.
FT   CONFLICT    215    216       TR -> SQ (IN REF. 1).
SQ   SEQUENCE   275 AA;  30772 MW;  B9BAC4BBCB91CE75 CRC64;

  Query Match             92.0%;  Score 1282;  DB 1;  Length 275;
  Best Local Similarity   93.1%;  Pred. No. 2e-110;
  Matches  228;  Conservative    4;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| |:|||||||| |||||
Db         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db        151 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||:||||||||||||||||||||||| |||||||||||||||||||||||||
Db        211 CAGNTRRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275


RESULT 4
MCT6_MOUSE
ID   MCT6_MOUSE     STANDARD;      PRT;   276 AA.
AC   P21845; Q61962;
DT   01-MAY-1991 (Rel. 18, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Mast cell protease 6 precursor (EC 3.4.21.59) (MMCP-6) (Tryptase).
GN   MCPT6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91139682; PubMed=1995638;
RA   Reynolds D.S., Gurley D.S., Austen K.F., Serafin W.E.;
RT   "Cloning of the cDNA and gene of mouse mast cell protease-6.
RT   Transcription by progenitor mast cells and mast cells of the
RT   connective tissue subclass.";
RL   J. Biol. Chem. 266:3847-3853(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Leaden X A1;
RX   MEDLINE=94023807; PubMed=8210998;
RA   Huang R., Abrink M., Gobl A.E., Nilsson G., Aveskogh M., Larsson L.G.,
RA   Nilsson K., Hellman L.;
RT   "Expression of a mast cell tryptase in the human monocytic cell lines
RT   U-937 and Mono Mac 6.";
RL   Scand. J. Immunol. 38:359-367(1993).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM SHORT).
RC   STRAIN=Leaden X A1;
RX   MEDLINE=95048582; PubMed=7959952;
RA   Huang R., Hellman L.T.;
RT   "Genes for mast-cell serine protease and their molecular evolution.";
RL   Immunogenetics 40:397-414(1994).
RN   [4]
RP   SEQUENCE OF 32-54.
```

```
RX   MEDLINE=90222202; PubMed=2326280;
RA   Reynolds D.S., Stevens R.L., Lane W.S., Carr M.H., Austen K.F.,
RA   Serafin W.E.;
RT   "Different mouse mast cell populations express various combinations
RT   of at least six distinct mast cell serine proteases.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:3230-3234(1990).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Long;
CC         IsoId=P21845-1; Sequence=Displayed;
CC       Name=Short;
CC         IsoId=P21845-2; Sequence=VSP_005376, VSP_005377;
CC         Note=Probably non functional;
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M57626; AAA39988.1; -.
DR   EMBL; M57625; AAA39987.1; -.
DR   EMBL; L31853; AAA39725.1; -.
DR   EMBL; X78542; CAA55288.1; -.
DR   PIR; A38654; A38654.
DR   PIR; I48685; I48685.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.025; -.
DR   MGD; MGI:96942; Mcpt6.
DR   GO; GO:0008201; F:heparin binding; IDA.
DR   GO; GO:0005515; F:protein binding; IPI.
DR   GO; GO:0030019; F:tryptase activity; IDA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;
KW   Alternative splicing.
FT   SIGNAL        1     21
FT   PROPEP       22     31       ACTIVATION PEPTIDE.
FT   CHAIN        32    276       MAST CELL PROTEASE 6.
FT   ACT_SITE     75     75       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    122    122       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    225    225       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     60     76       BY SIMILARITY.
FT   DISULFID    156    231       BY SIMILARITY.
FT   DISULFID    189    212       BY SIMILARITY.
FT   DISULFID    221    249       BY SIMILARITY.
FT   CARBOHYD    133    133       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    223    230       GDSGGPLV -> PFCIGDDI (in isoform Short).
FT                                /FTId=VSP_005376.
FT   VARSPLIC    231    276       Missing (in isoform Short).
FT                                /FTId=VSP_005377.
SQ   SEQUENCE   276 AA;  30927 MW;  525B2C9A04A72200 CRC64;

  Query Match             78.3%;  Score 1091;  DB 1;  Length 276;
  Best Local Similarity   78.2%;  Pred. No. 7e-93;
  Matches  190;  Conservative   19;  Mismatches   34;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||| ||  ||||||||||    ||:||||||||||||||||| |||| :|     ||||
Db         32 IVGGHEASESKWPWQVSLRFKLNYWIHFCGGSLIHPQWVLTAAHCVGPHIKSPQLFRVQL 91

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |||:||| |||| ::||:||| :|||: |||:|||||| || ||:|:| ::|||||||||
Db         92 REQYLYYGDQLLSLNRIVVHPHYYTAEGGADVALLELEVPVNVSTHIHPISLPPASETFP 151

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||  ||||||||:|||| ||||:|||||||||:|| :|| ||| | |||||  || | ||
Db        152 PGTSCWVTGWGDIDNDEPLPPPYPLKQVKVPIVENSLCDRKYHTGLYTGDDFPIVHDGML 211

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              |||||||||||||||||||||| |||||||||||||||||||:||||||||||||||| |
Db        212 CAGNTRRDSCQGDSGGPLVCKVKGTWLQAGVVSWGEGCAQPNKPGIYTRVTYYLDWIHRY 271

Qy        245 VPK 247
              ||:
Db        272 VPE 274


RESULT 5
TRYT_CANFA
ID   TRYT_CANFA     STANDARD;      PRT;   275 AA.
AC   P15944;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tryptase precursor (EC 3.4.21.59).
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89352460; PubMed=2504277;
RA   Vanderslice P., Craik C.S., Nadel J.A., Caughey G.H.;
RT   "Molecular cloning of dog mast cell tryptase and a related protease:
RT   structural evidence of a unique mode of serine protease activation.";
RL   Biochemistry 28:4148-4155(1989).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M24664; AAA30854.1; -.
DR   PIR; A32410; A32410.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.143; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen.
FT   SIGNAL        1     20       POTENTIAL.
FT   PROPEP       21     30       ACTIVATION PEPTIDE (BY SIMILARITY).
```

```
FT   CHAIN        31    275       TRYPTASE.
FT   ACT_SITE     74     74       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    121    121       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    224    224       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     59     75       BY SIMILARITY.
FT   DISULFID    155    230       BY SIMILARITY.
FT   DISULFID    188    211       BY SIMILARITY.
FT   DISULFID    220    248       BY SIMILARITY.
FT   CARBOHYD    132    132       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   275 AA;  30088 MW;  C3B869251F248D5B CRC64;

  Query Match             77.3%;  Score 1077;  DB 1;  Length 275;
  Best Local Similarity   77.1%;  Pred. No. 1.3e-91;
  Matches  189;  Conservative   21;  Mismatches   35;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||:||| ||||||||||: | || | ||||||||||||||| ||||:|     :||||
Db         31 IVGGREAPGSKWPWQVSLRLKGQYWRHICGGSLIHPQWVLTAAHCVGPNVVCPEEIRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||| ||||:||::|| :|| : ||||||||||:|| ||:||  |||||| :|||
Db         91 REQHLYYQDHLLPVNRIVMHPNYYTPENGADIALLELEDPVNVSAHVQPVTLPPALQTFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               | |||||||||| :   ||||||||||||||:|| :|| :||||  ||| |||||:|||
Db        151 TGTPCWVTGWGDVHSGTPLPPPFPLKQVKVPIVENSMCDVQYHLGLSTGDGVRIVREDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||:: |||||||||||||:| | ||||||||||||||||||||||||| ||||||| |
Db        211 CAGNSKSDSCQGDSGGPLVCRVRGVWLQAGVVSWGEGCAQPNRPGIYTRVAYYLDWIHQY 270

Qy        245 VPKKP 249
              |||:|
Db        271 VPKEP 275

RESULT 6
TRYT_MERUN
ID   TRYT_MERUN     STANDARD;      PRT;   270 AA.
AC   P50342;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Mast cell tryptase precursor (EC 3.4.21.59).
OS   Meriones unguiculatus (Mongolian jird) (Mongolian gerbil).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Gerbillinae;
OC   Meriones.
OX   NCBI_TaxID=10047;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MGS/SEA; TISSUE=Intestine;
RX   MEDLINE=95366971; PubMed=7639711;
RA   Murakumo Y., Ide H., Itoh H., Tomita M., Kobayashi T.,
RA   Maruyama H., Horii Y., Nawa Y.;
RT   "Cloning of the cDNA encoding mast cell tryptase of Mongolian gerbil,
RT   Meriones unguiculatus, and its preferential expression in the
RT   intestinal mucosa.";
RL   Biochem. J. 309:921-926(1995).
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D31789; BAA06598.1; -.
DR   PIR; S56160; S56160.
```

```
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.143; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Glycoprotein.
FT   SIGNAL        1     25       POTENTIAL.
FT   CHAIN        26    270       MAST CELL TRYPTASE.
FT   ACT_SITE     69     69       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    116    116       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    219    219       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     54     70       BY SIMILARITY.
FT   DISULFID    150    225       BY SIMILARITY.
FT   DISULFID    183    206       BY SIMILARITY.
FT   DISULFID    215    243       BY SIMILARITY.
FT   CARBOHYD     46     46       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    127    127       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    228    228       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   270 AA;  30166 MW;  1BE102DB86943401 CRC64;

  Query Match             77.2%;  Score 1075;  DB 1;  Length 270;
  Best Local Similarity   77.4%;  Pred. No. 2e-91;
  Matches  188;  Conservative   18;  Mismatches   37;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||||||| :||||||||| :  || ||||||||||||||||| |||| : |   :||||
Db         26 IVGGQEAPGNKWPWQVSLRANETYWRHFCGGSLIHPQWVLTAAHCVGPTIADPNKVRVQL 85

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |:|:||| | || ||||| || ||  | |||||||||: || :||||| |:|||||||||
Db         86 RKQYLYYHDHLLAVSRIITHPTFYATQNGADIALLELKNPVNISSHVHPVSLPPASETFP 145

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               |  |||||||::|||  ||||||||:|:||::|| :|| ||| | ||||:: |||||||
Db        146 SGTLCWVTGWGNIDNDVSLPPPFPLKEVQVPVVENQLCDLKYHKGVYTGDNIHIVRDDML 205

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||   |||||||||||||||||||||||||||||||| |||||||||||||||||| |
Db        206 CAGNEGHDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCALPNRPGIYTRVTYYLDWIHRY 265

Qy        245 VPK 247
              |||
Db        266 VPK 268

RESULT 7
MCT7_MOUSE
ID   MCT7_MOUSE     STANDARD;      PRT;   273 AA.
AC   Q02844;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Mast cell protease 7 precursor (EC 3.4.21.59) (MMCP-7) (Tryptase).
GN   MCPT7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2;
RX   MEDLINE=93087489; PubMed=1454796;
RA   McNeil H.P., Reynolds D.S., Schiller V., Ghildyal N., Gurley D.S.,
RA   Austen K.F., Stevens R.L.;
RT   "Isolation, characterization, and transcription of the gene encoding
RT   mouse mast cell protease 7.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:11174-11178(1992).
```

```
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM 2).
RC   STRAIN=C57BL/6J;
RX   MEDLINE=96162035; PubMed=8576265;
RA   Hunt J.E., Stevens R.L., Austen K.F., Zhang J., Xia Z.,
RA   Ghildyal N.;
RT   "Natural disruption of the mouse mast cell protease 7 gene in the
RT   C57BL/6 mouse.";
RL   J. Biol. Chem. 271:2851-2855(1996).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC         Comment=The alternative splicing event is due to a G to A point
CC         mutation at the exon 2/intron 2 splice site and causes loss of
CC         protein expression. The alternatively spliced transcript is only
CC         found in C57BL/6 mouse;
CC       Name=1;
CC         IsoId=Q02844-1; Sequence=Displayed;
CC       Name=2; Synonyms=Truncated;
CC         IsoId=Q02844-2; Sequence=VSP_005378, VSP_005379;
CC   -!- DEVELOPMENTAL STAGE: Is not expressed in mature serosal or mucosal
CC       mast cells and is expressed only transiently at an early stage of
CC       in vitro mast cell differentiation.
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; L00654; AAA39993.1; -.
DR   EMBL; L00653; AAA39992.1; -.
DR   EMBL; U42405; AAA97874.1; -.
DR   EMBL; U42406; AAA97875.1; -.
DR   PIR; A47246; A47246.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.026; -.
DR   MGD; MGI:96943; Mcpt7.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Zymogen; Alternative splicing;
KW   Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     28       ACTIVATION PEPTIDE (POTENTIAL).
FT   CHAIN        29    273       MAST CELL PROTEASE 7.
FT   ACT_SITE     72     72       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    119    119       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    222    222       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     57     73       BY SIMILARITY.
FT   DISULFID    153    228       BY SIMILARITY.
FT   DISULFID    186    209       BY SIMILARITY.
FT   DISULFID    218    246       BY SIMILARITY.
FT   CARBOHYD     49     49       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    130    130       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     44     46       VSL -> GCC (in isoform 2).
FT                                /FTId=VSP_005378.
FT   VARSPLIC     47    273       Missing (in isoform 2).
FT                                /FTId=VSP_005379.
SQ   SEQUENCE   273 AA;  30337 MW;  50ECB4765294205E CRC64;
```

```
  Query Match             75.7%;  Score 1054;  DB 1;  Length 273;
  Best Local Similarity   76.5%;  Pred. No. 1.7e-89;
  Matches  186;  Conservative   17;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||||||  :||||||||| :  |||||||||||||||||||| |||||| |   :||||
Db         29 IVGGQEAHGNKWPWQVSLRANDTYWMHFCGGSLIHPQWVLTAAHCVGPDVADPNKVRVQL 88

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |:|:||| | |: ||:|| || ||  | |||||||:|  || :| :|| | |||||||||
Db         89 RKQYLYYHDHLMTVSQIITHPDFYIVQDGADIALLKLTNPVNISDYVHPVPLPPASETFP 148

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               |  |||||||::||   ||||||||:|:|||:|||:|| ||| |  |||:| |||||||
Db        149 SGTLCWVTGWGNIDNGVNLPPPFPLKEVQVPIIENHLCDLKYHKGLITGDNVHIVRDDML 208

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||   |||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db        209 CAGNEGHDSCQGDSGGPLVCKVEDTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 268

Qy        245 VPK 247
              |||
Db        269 VPK 271


RESULT 8
MCT6_RAT
ID   MCT6_RAT       STANDARD;      PRT;   274 AA.
AC   P50343; P97593;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Mast cell protease 6 precursor (EC 3.4.21.59) (RMCP-6) (Tryptase).
GN   MCPT6 OR MCP6.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Peritoneal mast cells;
RX   MEDLINE=96015171; PubMed=8537314;
RA   Ide H., Itoh H., Tomita M., Murakumo Y., Kobayashi T.,
RA   Maruyama H., Osada Y., Nawa Y.;
RT   "cDNA sequencing and expression of rat mast cell tryptase.";
RL   J. Biochem. 118:210-215(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Peritoneal mast cells;
RX   MEDLINE=97149430; PubMed=8996238;
RA   Lutzelschwab C., Pejler G., Aveskogh M., Hellman L.;
RT   "Secretory granule proteases in rat mast cells. Cloning of 10
RT   different serine proteases and a carboxypeptidase A from various rat
RT   mast cell populations.";
RL   J. Exp. Med. 185:13-29(1997).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D38455; BAA07486.1; -.
DR   EMBL; U67909; AAB48262.1; -.
```

```
DR   PIR; JC4171; JC4171.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.025; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Glycoprotein; Multigene family;
KW   Signal.
FT   SIGNAL        1     19       POTENTIAL.
FT   PROPEP       20     29       ACTIVATION PEPTIDE.
FT   CHAIN        30    274       MAST CELL PROTEASE 6.
FT   ACT_SITE     73     73       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    120    120       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    223    223       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     58     74       BY SIMILARITY.
FT   DISULFID    154    229       BY SIMILARITY.
FT   DISULFID    187    210       BY SIMILARITY.
FT   DISULFID    219    247       BY SIMILARITY.
FT   CARBOHYD    104    104       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    131    131       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    128    128       I -> N (IN REF. 2).
FT   CONFLICT    139    139       I -> T (IN REF. 2).
SQ   SEQUENCE   274 AA;  30508 MW;  DF84D55668CA1A25 CRC64;

  Query Match             75.4%;  Score 1051;  DB 1;  Length 274;
  Best Local Similarity   75.8%;  Pred. No. 3.3e-89;
  Matches  185;  Conservative   21;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||:||  ||||||||||    :||||||||||||||||||| |||  :|     ||||
Db         30 IVGGREASESKWPWQVSLRFKFSFWMHFCGGSLIHPQWVLTAAHCVGLHIKSPELFRVQL 89

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |||:||| |||| |:| :||| :|| : |||||||||| || ||:|:| ::|||||||||
Db         90 REQYLYYADQLLTVNRTVVHPHYYTVEDGADIALLELEIPVNVSTHIHPISLPPASETFP 149

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               |  ||||||||:|:|| | ||:|||||||||:|| :|| ||| | |||||| ||:| ||
Db        150 SGTSCWVTGWGDIDSDEPLLPPYPLKQVKVPIVENSLCDRKYHTGLYTGDDVPIVQDGML 209

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              |||||| ||||||||||||||| ||||||||||||||||: ||||||||||||||||| |
Db        210 CAGNTRSDSCQGDSGGPLVCKVKGTWLQAGVVSWGEGCAEANRPGIYTRVTYYLDWIHRY 269

Qy        245 VPKK 248
              ||::
Db        270 VPQR 273


RESULT 9
MCT7_RAT
ID   MCT7_RAT       STANDARD;      PRT;   273 AA.
AC   P27435; P27436;
DT   01-AUG-1992 (Rel. 23, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Mast cell protease 7 precursor (EC 3.4.21.59) (RMCP-7) (Tryptase,
DE   skin).
GN   MCPT7 OR MCP7.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley;
RX   MEDLINE=97149430; PubMed=8996238;
RA   Lutzelschwab C., Pejler G., Aveskogh M., Hellman L.;
```

```
RT   "Secretory granule proteases in rat mast cells. Cloning of 10
RT   different serine proteases and a carboxypeptidase A from various rat
RT   mast cell populations.";
RL   J. Exp. Med. 185:13-29(1997).
RN   [2]
RP   SEQUENCE OF 29-53.
RC   STRAIN=Sprague-Dawley; TISSUE=Skin;
RX   MEDLINE=91242400; PubMed=2036367;
RA   Braganza V.J., Simmons W.H.;
RT   "Tryptase from rat skin: purification and properties.";
RL   Biochemistry 30:4997-5007(1991).
RN   [3]
RP   SEQUENCE OF 29-51.
RC   TISSUE=Breast carcinoma;
RX   MEDLINE=92231826; PubMed=1314562;
RA   Eto I., Grubbs C.J.;
RT   "Separation, purification and N-terminal sequence analysis of a novel
RT   leupeptin-sensitive serine endopeptidase present in chemically
RT   induced rat mammary tumour.";
RL   Biochem. J. 283:209-216(1992).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
CC   -!- TISSUE SPECIFICITY: Mast cells.
CC   -!- PTM: Glycosylated (Probable).
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; U67910; AAB48263.1; -.
DR   PIR; A23698; A23698.
DR   PIR; S21275; S21275.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.026; -.
DR   MEROPS; S01.143; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal;
KW   Multigene family.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     28       ACTIVATION PEPTIDE.
FT   CHAIN        29    273       MAST CELL PROTEASE 7.
FT   ACT_SITE     72     72       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    119    119       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    222    222       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     57     73       BY SIMILARITY.
FT   DISULFID    153    228       BY SIMILARITY.
FT   DISULFID    186    209       BY SIMILARITY.
FT   DISULFID    218    246       BY SIMILARITY.
FT   CARBOHYD     49     49       N-LINKED (GLCNAC. . .) (PROBABLE).
FT   CONFLICT     42     42       W -> V (IN REF. 3).
FT   CONFLICT     49     51       NDT -> WLP (IN REF. 3).
SQ   SEQUENCE   273 AA;  30400 MW;  65A5ED4D279FB284 CRC64;

  Query Match             74.4%;  Score 1037;  DB 1;  Length 273;
```

```
  Best Local Similarity   76.1%;  Pred. No. 6.3e-88;
  Matches  185;  Conservative   21;  Mismatches   37;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||||||  :||||||||||:  |||||||||||||||||||| ||||:  |   |||||
Db         29 IVGGQEASGNKWPWQVSLRVNDTYWMHFCGGSLIHPQWVLTAAHCVGPNKADPNKLRVQL 88

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |:|:||| | || ||:|| || || || |||||||:|  || ::|:||||:|||||||||
Db         89 RKQYLYYHDHLLTVSQIISHPDFYIAQDGADIALLKLTNPVNITSNVHTVSLPPASETFP 148

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               |  |||||||:::||  |||||||::|:|||:|| :|| ||| |  |||:| |||||||
Db        149 SGTLCWVTGWGNINNDVSLPPPFPLEEVQVPIVENRLCDLKYHKGLNTGDNVHIVRDDML 208

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||   |||||||||||||||  |||||||||||||||||||||||||||||||||: |
Db        209 CAGNEGHDSCQGDSGGPLVCKVEDTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIYRY 268

Qy        245 VPK 247
              |||
Db        269 VPK 271


RESULT 10
TRYT_SHEEP
ID   TRYT_SHEEP     STANDARD;      PRT;   273 AA.
AC   Q9XSM2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tryptase 2 precursor (EC 3.4.21.59).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Abomasum;
RX   MEDLINE=20308142; PubMed=10848900;
RA   Pemberton A.D., McAleese S.M., Huntley J.F., Collie D.D.S.,
RA   Scudamore C.L., McEuen A.R., Walls A.F., Miller H.R.P.;
RT   "cDNA sequence of two sheep mast cell tryptases and the differential
RT   expression of tryptase and sheep mast cell proteinase-1 in lung,
RT   dermis and gastrointestinal tract.";
RL   Clin. Exp. Allergy 30:818-832(2000).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; Y18224; CAB41989.1; -.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.143; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
```

```
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     28       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   CHAIN        29    273       TRYPTASE 2.
FT   ACT_SITE     72     72       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    119    119       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    222    222       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     57     73       BY SIMILARITY.
FT   DISULFID    153    228       BY SIMILARITY.
FT   DISULFID    186    209       BY SIMILARITY.
FT   DISULFID    218    246       BY SIMILARITY.
FT   CARBOHYD    231    231       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   273 AA;  30288 MW;  DE9BA79218C3E67D CRC64;

  Query Match             73.9%;  Score 1030;  DB 1;  Length 273;
  Best Local Similarity   73.1%;  Pred. No. 2.8e-87;
  Matches  179;  Conservative   27;  Mismatches   39;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |:||:||| |:|||||||||   || | ||||||||||||||| |:||:::: :  ||||
Db         29 IIGGKEAPGSRWPWQVSLRVRDQYWRHQCGGSLIHPQWVLTAAHCIGPELQEPSDFRVQL 88

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |||||||||:|||:||:| || :|  : |||||||:||||| :| ||  |||||||||||
Db         89 REQHLYYQDRLLPISRVIPHPHYYMVENGADIALLQLEEPVSISRHVQPVTLPPASETFP 148

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              |   |||||||||||   ||||:|||||||||:|| :|| ||| |  |   | ||::| |
Db        149 PESQCWVTGWGDVDNGRPLPPPYPLKQVKVPIVENSVCDWKYHSGLSTDYSVPIVQEDNL 208

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              |||:  |||||||||||||||||||||||||||||:|||:|||||||||:| |||||| |
Db        209 CAGDGGRDSCQGDSGGPLVCKVNGTWLQAGVVSWGDGCAKPNRPGIYTRITSYLDWIHQY 268

Qy        245 VPKKP 249
              ||::|
Db        269 VPQEP 273


RESULT 11
TRYT_PIG
ID   TRYT_PIG       STANDARD;      PRT;   275 AA.
AC   Q9N2D1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Tryptase precursor (EC 3.4.21.59).
GN   MCT7.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=20285343; PubMed=10824103;
RA   Chen Y., Shiota M., Ohuchi M., Towatari T., Tashiro J., Murakami M.,
RA   Yano M., Yang B., Kido H.;
RT   "Mast cell tryptase from pig lungs triggers infection by pneumotropic
RT   Sendai and influenza A viruses. Purification and characterization.";
RL   Eur. J. Biochem. 267:3189-3197(2000).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation.
```

```
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AB038652; BAA93614.1; -.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.143; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen.
FT   SIGNAL        1     20       POTENTIAL.
FT   PROPEP       21     30       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   CHAIN        31    275       TRYPTASE.
FT   ACT_SITE     74     74       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    121    121       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    224    224       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     59     75       BY SIMILARITY.
FT   DISULFID    155    230       BY SIMILARITY.
FT   DISULFID    188    211       BY SIMILARITY.
FT   DISULFID    220    248       BY SIMILARITY.
FT   CARBOHYD    132    132       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    233    233       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   275 AA;  30439 MW;  ACC582647FCCB973 CRC64;

  Query Match             72.2%;  Score 1006;  DB 1;  Length 275;
  Best Local Similarity   74.2%;  Pred. No. 4.5e-85;
  Matches  181;  Conservative   20;  Mismatches   43;  Indels    0;  Gaps    0;

Qy        5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
            ||||:|||  |||||||||    || ||||||||||||||||| | ||:  |   :||||
Db       31 IVGGKEAPGHKWPWQVSLRCLDQYWKHFCGGSLIHPQWVLTAAHCFGPEKADPLYIRVQL 90

Qy       65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
             ||||||||:|| |||||||| :|    ||||||||||:|| :||||  |||||||||||
Db       91 GEQHLYYQDRLLLVSRIIVHPNYYDEVNGADIALLELEDPVNLSSHVQPVTLPPASETFP 150

Qy      125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
             |  ||||||||| :   ||||:|||||:|||:||  || :||||  |||:: |||||||
Db      151 KGTRCWVTGWGDVHSGWPLPPPYPLKQVRVPIVENSECDMQYHLGLSTGDNIPIVRDDML 210

Qy      185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
            |||:   |||||||||||||:|||||||||||||||||| |||||||||||:||||||
Db      211 CAGSEGHDSCQGDSGGPLVCRVNGTWLQAGVVSWGEGCALPNRPGIYTRVTHYLDWIHQC 270

Qy      245 VPKK 248
            :|::
Db      271 IPRE 274


RESULT 12
TRYD_HUMAN
ID   TRYD_HUMAN     STANDARD;      PRT;   235 AA.
AC   Q9BZJ3; O95824; Q8TDI6; Q96L36; Q96RZ5; Q9H2Y6; Q9UQI8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   10-OCT-2003 (Rel. 42, Last annotation update)
DE   Tryptase delta precursor (EC 3.4.21.59) (Delta tryptase) (Mast cell
DE   mMCP-7-like) (HmMCP-3-like tryptase III).
GN   TPSD1.
OS   Homo sapiens (Human).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND VARIANTS ARG-15 AND MET-76.
RX   MEDLINE=99121069; PubMed=9920877;
RA   Pallaoro M., Fejzo M.S., Shayesteh L., Blount J.L., Caughey G.H.;
RT   "Characterization of genes encoding known and novel human mast cell
RT   tryptases on chromosome 16p13.3.";
RL   J. Biol. Chem. 274:3355-3362(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21101554; PubMed=11174199;
RA   Min H.-K., Kambe N., Schwartz L.B.;
RT   "Human mouse mast cell protease 7-like tryptase genes are
RT   pseudogenes.";
RL   J. Allergy Clin. Immunol. 107:315-321(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21096910; PubMed=11157797;
RA   Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA   Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA   Higgs D.R.;
RT   "Sequence, structure and pathology of the fully annotated terminal 2
RT   Mb of the short arm of human chromosome 16.";
RL   Hum. Mol. Genet. 10:339-352(2001).
RN   [4]
RP   SEQUENCE OF 33-235 FROM N.A. (ISOFORMS 1 AND 2), ENZYMATIC ACTIVITY,
RP   TISSUE SPECIFICITY, AND VARIANT MET-76.
RX   MEDLINE=22278650; PubMed=12391231;
RA   Wang H.-W., McNeil H.P., Husain A., Liu K., Tedla N., Thomas P.S.,
RA   Raftery M., King G.C., Cai Z.Y., Hunt J.E.;
RT   "Delta tryptase is expressed in multiple human tissues, and a
RT   recombinant form has proteolytic activity.";
RL   J. Immunol. 169:5145-5152(2002).
CC   -!- FUNCTION: Tryptase is the major neutral protease present in mast
CC       cells and is secreted upon the coupled activation-degranulation
CC       response of this cell type (By similarity).
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-, Lys-|-, but
CC       with more restricted specificity than trypsin.
CC   -!- SUBUNIT: Homotetramer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Released from the secretory granules upon
CC       mast cell activation (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=1;
CC         IsoId=Q9BZJ3-1; Sequence=Displayed;
CC       Name=2;
CC         IsoId=Q9BZJ3-2; Sequence=VSP_008319;
CC   -!- TISSUE SPECIFICITY: Expressed in colon, lung, heart and synovial
CC       tissue. May be specific to mast cells.
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   -!- CAUTION: Although Ref.2 reported this as a pseudogene, Ref.4
CC       showed it is expressed and has proteolytic activity when expressed
CC       in bacterial cells.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF098327; AAD17845.1; ALT_INIT.
DR   EMBL; AF099147; AAD17861.1; -.
DR   EMBL; AF318074; AAK12909.1; -.
DR   EMBL; AE006466; AAK61272.1; ALT_INIT.
DR   EMBL; AF206664; AAG35694.1; -.
DR   EMBL; AF421357; AAL86695.1; -.
DR   EMBL; AY055427; AAL17874.1; -.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.054; -.
```

```
DR   Genew; HGNC:14118; TPSD1.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Zymogen; Glycoprotein;
KW   Alternative splicing; Polymorphism.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     30       ACTIVATION PEPTIDE (BY SIMILARITY).
FT   CHAIN        31    235       TRYPTASE DELTA.
FT   ACT_SITE     74     74       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    121    121       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    224    224       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     59     75       BY SIMILARITY.
FT   DISULFID    155    230       BY SIMILARITY.
FT   DISULFID    188    211       BY SIMILARITY.
FT   CARBOHYD    132    132       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC     79     87       Missing (in isoform 2).
FT                                /FTId=VSP_008319.
FT   VARIANT      15     15       P -> R (in dbSNP:3865205).
FT                                /FTId=VAR_016870.
FT   VARIANT      18     18       V -> A (in dbSNP:1800984).
FT                                /FTId=VAR_016871.
FT   VARIANT      76     76       V -> M (in dbSNP:3993987).
FT                                /FTId=VAR_016872.
FT   CONFLICT     18     18       V -> G (IN REF. 1; AAD17861).
FT   CONFLICT    171    171       P -> S (IN REF. 4; AAL17874).
SQ   SEQUENCE   235 AA;  25816 MW;  B1FFB6C2A8006B22 CRC64;

  Query Match             70.9%;  Score 988;  DB 1;  Length 235;
  Best Local Similarity   84.4%;  Pred. No. 1.7e-83;
  Matches  173;  Conservative   15;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||||||||||||||||||| |||||||||||||||||||||| || ||:||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVRGPYWMHFCGGSLIHPQWVLTAAHCVEPDIKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||  | ||||||||||||| :|||:|||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSHIHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              |||||||||||||||:  ||||:|||:|:||::|||:|:|:|| | :||    :|||||||
Db        151 PGMPCWVTGWGDVDNNVHLPPPYPLKEVEVPVVENHLCNAEYHTGLHTGHSFQIVRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGT 209
              |||:   ||||||||||||||||||
Db        211 CAGSENHDSCQGDSGGPLVCKVNGT 235


RESULT 13
TRYM_CANFA
ID   TRYM_CANFA     STANDARD;      PRT;   269 AA.
AC   P19236;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Mastocytoma protease precursor (EC 3.4.21.-).
OS   Canis familiaris (Dog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX   NCBI_TaxID=9615;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89352460; PubMed=2504277;
RA   Vanderslice P., Craik C.S., Nadel J.A., Caughey G.H.;
RT   "Molecular cloning of dog mast cell tryptase and a related protease:
RT   structural evidence of a unique mode of serine protease activation.";
RL   Biochemistry 28:4148-4155(1989).
CC   -!- FUNCTION: Mast cell protease.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; M24665; AAA30855.1; -.
DR   PIR; B32410; B32410.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.145; -.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
KW   Hydrolase; Serine protease; Signal; Glycoprotein.
FT   SIGNAL        1     19
FT   CHAIN        20    269       MASTOCYTOMA PROTEASE.
FT   ACT_SITE     66     66       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    116    116       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    217    217       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     51     67       BY SIMILARITY.
FT   DISULFID    150    223       BY SIMILARITY.
FT   DISULFID    183    204       BY SIMILARITY.
FT   DISULFID    213    241       BY SIMILARITY.
FT   CARBOHYD     95     95       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    106    106       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   269 AA;  29824 MW;  B4A4AF7E4E2FD4E5 CRC64;

  Query Match             49.1%;  Score 684.5;  DB 1;  Length 269;
  Best Local Similarity   52.6%;  Pred. No. 1.4e-55;
  Matches  132;  Conservative   32;  Mismatches   78;  Indels    9;  Gaps    4;

Qy          5 IVGGQEAPRSKWPWQVSLRVHG---PYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALR 61
              |||| : |  ::||||||| ||     | | ||||||||||||||| ||  :  :  | ||
Db         20 IVGGCKVPARRYPWQVSLRFHGMGSGQWQHICGGSLIHPQWVLTAAHCVELEGLEAATLR 79

Qy         62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIG---ADIALLELEEPVKVSSHVHTVTLPP 118
              ||: :  ||  |||  |: || || |  :  |   ||||||:|| |: :|  |: |:|| 
Db         80 VQVGQLRLYDHDQLCNVTEIIRHPNFNMSWYGWDTADIALLKLEAPLTLSEDVNLVSLPS 139

Qy        119 ASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRI 178
               |   |||| ||||||||: :   |||| |::|:|||: |  |:  |        |  :
Db        140 PSLIVPPGMLCWVTGWGDIADHTPLPPPYHLQEVEVPIVGNRECNCHYQ--TILEQDDEV 197

Qy        179 VRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYL 238
              :: ||||||:   |||| ||||||||:   ||:| |||||| ||   | ||:| ||| |:
Db        198 IKQDMLCAGSEGHDSCQMDSGGPLVCRWKCTWIQVGVVSWGYGCGY-NLPGVYARVTSYV 256

Qy        239 DWIHHYVPKKP 249
               ||| ::|  |
Db        257 SWIHQHIPLSP 267


RESULT 14
TRYG_MOUSE
ID   TRYG_MOUSE     STANDARD;      PRT;   311 AA.
AC   Q9QUL7;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-MAR-2004 (Rel. 43, Last annotation update)
DE   Tryptase gamma precursor (EC 3.4.21.-) (Transmembrane tryptase).
GN   TPSG1 OR TMT.
```

```
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/Sv, and BALB/c;
RX   MEDLINE=99452974; PubMed=10521469;
RA   Wong G.W., Tang Y., Feyfant E., Sali A., Li L., Li Y., Huang C.,
RA   Friend D.S., Krilis S.A., Stevens R.L.;
RT   "Identification of a new member of the tryptase family of mouse and
RT   human mast cell proteases which possesses a novel COOH-terminal
RT   hydrophobic extension.";
RL   J. Biol. Chem. 274:30784-30793(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Colon;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Membrane-anchored (Potential).
CC   -!- TISSUE SPECIFICITY: Expressed in many tissues.
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF175760; AAF03698.1; -.
DR   EMBL; AF175523; AAF03696.1; -.
DR   EMBL; BC052325; AAH52325.1; -.
DR   HSSP; P20231; 1AAO.
DR   MEROPS; S01.028; -.
DR   MGD; MGI:1349391; Tpsg1.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
KW   Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;
KW   Transmembrane.
FT   SIGNAL        1     16       POTENTIAL.
FT   CHAIN        17     28       TRYPTASE GAMMA LIGHT CHAIN.
FT   CHAIN        30    311       TRYPTASE GAMMA HEAVY CHAIN.
FT   TRANSMEM    277    297       POTENTIAL.
FT   ACT_SITE     70     70       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    117    117       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    214    214       CHARGE RELAY SYSTEM (BY SIMILARITY).
```

```
FT   DISULFID     18    137       INTERCHAIN (POTENTIAL).
FT   DISULFID     55     71       BY SIMILARITY.
FT   DISULFID    151    220       BY SIMILARITY.
FT   DISULFID    184    202       BY SIMILARITY.
FT   DISULFID    210    238       BY SIMILARITY.
FT   CARBOHYD     77     77       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    192    192       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   311 AA;  32656 MW;  7FC9D6EF6A2A8808 CRC64;

  Query Match             43.5%;  Score 605.5;  DB 1;  Length 311;
  Best Local Similarity   48.6%;  Pred. No. 3.1e-48;
  Matches  121;  Conservative   31;  Mismatches   78;  Indels   19;  Gaps    6;

Qy          4 RIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQ 63
              |||||  ||   |||| |||:|    :| |||||: |:|||||| |    |   :  :|
Db         29 RIVGGHAAPAGTWPWQASLRLHK---VHVCGGSLLSPEWVLTAAHCFSGSVNS-SDYQVH 84

Qy         64 LREQHLYYQDQLLPVSRIIVHPQFYTAQIG-----ADIALLELEEPVKVSSHVHTVTLPP 118
              | |  :        | |||:    ||   |      ||||::|  || :|| |  | ||
Db         85 LGELTVTLSPHFSTVKRIIM----YTGSPGPPGSSGDIALVQLSSPVALSSQVQPVCLPE 140

Qy        119 ASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRI 178
              ||  | ||| |||||||     | | ||: |:: || :::   |       ||   :  :
Db        141 ASADFYPGMQCWVTGWGYTGEGEPLKPPYNLQEAKVSVVDVKTCSQ-----AYNSPNGSL 195

Qy        179 VRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYL 238
              :: |||||     |:|| ||||||||:| ||| ||||||||||| :|:|||:| ||| |:
Db        196 IQPDMLCARGP-GDACQDDSGGPLVCQVAGTWQQAGVVSWGEGCGRPDRPGVYARVTAYV 254

Qy        239 DWIHHYVPK 247
              :||||::|:
Db        255 NWIHHHIPE 263


RESULT 15
TRYG_HUMAN
ID   TRYG_HUMAN     STANDARD;      PRT;   321 AA.
AC   Q9NRR2; Q9C015; Q9NRQ8; Q9UBB2;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Tryptase gamma precursor (EC 3.4.21.-) (Transmembrane tryptase).
GN   TPSG1 OR TMT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (VARIANTS GAMMA-1 AND GAMMA-2).
RX   MEDLINE=20302813; PubMed=10843716;
RA   Caughey G.H., Raymond W.W., Blount J.L., Hau L.W., Pallaoro M.,
RA   Wolters P.J., Verghese G.M.;
RT   "Characterization of human gamma-tryptases, novel members of the
RT   chromosome 16p mast cell tryptase and prostasin gene families.";
RL   J. Immunol. 164:6566-6575(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99452974; PubMed=10521469;
RA   Wong G.W., Tang Y., Feyfant E., Sali A., Li L., Li Y., Huang C.,
RA   Friend D.S., Krilis S.A., Stevens R.L.;
RT   "Identification of a new member of the tryptase family of mouse and
RT   human mast cell proteases which possesses a novel COOH-terminal
RT   hydrophobic extension.";
RL   J. Biol. Chem. 274:30784-30793(1999).
RN   [3]
RP   SEQUENCE OF 220-321 FROM N.A.
RA   Mittman S., Agnew W.S.;
RT   "Organization and alternative splicing of CACNA1H.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Membrane-anchored (Potential).
CC   -!- TISSUE SPECIFICITY: Expressed in many tissues.
CC   -!- POLYMORPHISM: There are two alleles; gamma-I and gamma-II which
CC       differ by 5 residues.
```

```
CC   -!- SIMILARITY: Belongs to peptidase family S1. Tryptase subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AF191031; AAF76457.1; -.
DR   EMBL; AF195508; AAF76458.1; -.
DR   EMBL; AF175759; AAF03697.1; -.
DR   EMBL; AF175522; AAF03695.1; -.
DR   EMBL; AF223563; AAG48852.2; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.028; -.
DR   Genew; HGNC:14134; TPSG1.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; FALSE_NEG.
KW   Hydrolase; Serine protease; Signal; Glycoprotein; Zymogen;
KW   Transmembrane; Polymorphism.
FT   SIGNAL        1     19       POTENTIAL.
FT   CHAIN        20     36       TRYPTASE GAMMA LIGHT CHAIN.
FT   CHAIN        38    321       TRYPTASE GAMMA HEAVY CHAIN.
FT   TRANSMEM    284    304       POTENTIAL.
FT   ACT_SITE     78     78       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    125    125       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   ACT_SITE    222    222       CHARGE RELAY SYSTEM (BY SIMILARITY).
FT   DISULFID     26    145       INTERCHAIN (POTENTIAL).
FT   DISULFID     63     79       BY SIMILARITY.
FT   DISULFID    159    228       BY SIMILARITY.
FT   DISULFID    192    210       BY SIMILARITY.
FT   DISULFID    218    246       BY SIMILARITY.
FT   CARBOHYD     85     85       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      60     60       M -> V (IN GAMMA-II).
FT                                /FTId=VAR_012097.
FT   VARIANT     126    126       I -> M (IN GAMMA-II).
FT                                /FTId=VAR_012098.
FT   VARIANT     132    132       S -> T (IN GAMMA-II).
FT                                /FTId=VAR_012099.
FT   VARIANT     204    204       L -> I (IN GAMMA-II).
FT                                /FTId=VAR_012100.
FT   VARIANT     288    288       L -> F (IN GAMMA-II).
FT                                /FTId=VAR_012101.
FT   CONFLICT    160    160       W -> S (IN REF. 1).
SQ   SEQUENCE   321 AA;  33827 MW;  FFF7B06E3C4A962D CRC64;

  Query Match             42.6%;  Score 593.5;  DB 1;  Length 321;
  Best Local Similarity   48.8%;  Pred. No. 4e-47;
  Matches  119;  Conservative   32;  Mismatches   80;  Indels   13;  Gaps    6;

Qy          4 RIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQ 63
              |||||  ||   |||| |||:     || |||||: |||||||| |    :   :  :|
Db         37 RIVGGHAAPAGAWPWQASLRLRR---MHVCGGSLLSPQWVLTAAHCFSGSLNS-SDYQVH 92

Qy         64 LREQHLYYQDQLLPVSRIIVHPQFYTAQIG--ADIALLELEEPVKVSSHVHTVTLPPASE 121
              | |  :        | :||:|    : | |   ||||:||  || :|| :  | || ||:
Db         93 LGELEITLSPHFSTVRQIILHSS-PSGQPGTSGDIALVELSVPVTLSSRILPVCLPEASD 151

Qy        122 TFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRD 181
               | ||: |||||||     | ||||: |::||| :::   |        | |    |::
Db        152 DFCPGIRCWVTGWGYTREGEPLPPPYSLREVKVSVVDTETCRRD-----YPGPGGSILQP 206

Qy        182 DMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWI 241
              |||||     |:|| ||||||||:||| |:|||:||||||| :|||||:||||  |::||
Db        207 DMLCARGP-GDACQDDSGGPLVCQVNGAWVQAGIVSWGEGCGRPNRPGVYTRVPAYVNWI 265

Qy        242 HHYV 245
                ::
Db        266 RRHI 269

Search completed: July 22, 2004, 15:23:30
Job time : 54 secs
```

```
                     GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          July 22, 2004, 13:24:46 ; Search time 98 Seconds
                                           (without alignments)
                                           717.901 Million cell updates/sec

Title:           US-09-598-982-21
Perfect score:   1393
Sequence:        1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters:      1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        A_Geneseq_29Jan04:*
                1:  geneseqp1980s:*
                2:  geneseqp1990s:*
                3:  geneseqp2000s:*
                4:  geneseqp2001s:*
                5:  geneseqp2002s:*
                6:  geneseqp2003as:*
                7:  geneseqp2003bs:*
                8:  geneseqp2004s:*

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.

                              SUMMARIES

                  %
Result           Query
  No.    Score   Match Length DB  ID                    Description
---------------------------------------------------------------------
    1     1393  100.0    249  5  AAU12009              Aau12009 Human bet
    2     1388   99.6    249  5  AAU12017              Aau12017 Human bet
    3     1387   99.6    249  3  AAY55011              Aay55011 Human bet
    4     1387   99.6    249  5  AAU12007              Aau12007 Recombina
    5     1384   99.4    249  5  AAU12011              Aau12011 Human bet
    6     1384   99.4    249  5  AAU12012              Aau12012 Human bet
    7     1382   99.2    249  5  AAU12006              Aau12006 Human bet
    8     1379   99.0    249  5  AAU12020              Aau12020 Human bet
    9     1379   99.0    249  5  AAU12010              Aau12010 Human bet
   10     1379   99.0    249  5  AAU12019              Aau12019 Human bet
   11     1374   98.6    245  5  AAU12013              Aau12013 Human bet
   12     1374   98.6    249  5  AAU12018              Aau12018 Human bet
   13     1369   98.3    245  5  AAU12021              Aau12021 Human bet
   14     1368   98.2    245  5  AAU12008              Aau12008 Recombina
   15     1368   98.2    274  2  AAW64240              Aaw64240 Human mas
   16     1368   98.2    274  2  AAW63175              Aaw63175 Human mas
   17     1365   98.0    245  5  AAU12016              Aau12016 Human bet
   18     1365   98.0    245  5  AAU12015              Aau12015 Human bet
   19     1363   97.8    245  3  AAY55010              Aay55010 Human bet
   20     1363   97.8    245  5  AAU12005              Aau12005 Human mat
   21     1363   97.8    273  2  AAW64238              Aaw64238 Human mas
   22     1363   97.8    273  2  AAW63174              Aaw63174 Human mas
   23     1361   97.7    244  2  AAY25925              Aay25925 Human lun
   24     1360   97.6    245  5  AAU12014              Aau12014 Human bet
   25     1360   97.6    245  5  AAU12023              Aau12023 Human bet
```

```
   26     1360   97.6    245  5  AAU12024              Aau12024 Human bet
   27     1355   97.3    245  5  AAU12022              Aau12022 Human bet
   28     1344   96.5    267  2  AAW64241              Aaw64241 Human mas
   29     1344   96.5    267  2  AAW63176              Aaw63176 Human mas
   30     1344   96.5    267  5  AAU84360              Aau84360 Protein T
   31     1339   96.1    275  7  ADE56171              Ade56171 Human Pro
   32     1329   95.4    691  5  AAE14348              Aae14348 Human pro
   33     1282   92.0    275  7  ADE62892              Ade62892 Human Pro
   34     1274   91.5    275  2  AAW64237              Aaw64237 Human mas
   35     1274   91.5    275  2  AAW63173              Aaw63173 Human mas
   36     1091   78.3    276  2  AAW64242              Aaw64242 Murine ma
   37     1091   78.3    276  2  AAW63172              Aaw63172 Mouse mas
   38     1077   77.3    276  5  ABB05437              Abb05437 Mouse mas
   39     1075   77.2    270  2  AAW64239              Aaw64239 Gerbil ho
   40     1054   75.7    273  2  AAW64233              Aaw64233 Murine mM
   41     1054   75.7    310  5  AAE14343              Aae14343 Human pro
   42     1051   75.4    274  2  AAW64234              Aaw64234 Rat homol
   43     1051   75.4    274  7  ADE56169              Ade56169 Rat Prote
   44     1037   74.4    273  2  AAW64235              Aaw64235 Rat homol
   45     1037   74.4    273  7  ADD47563              Add47563 Rat Prote

                              ALIGNMENTS

RESULT 1
AAU12009
ID   AAU12009 standard; protein; 249 AA.
XX
AC   AAU12009;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human beta-II tryptase active site mutant H44A #1.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20775.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 8; Page 84-85; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
```

```
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAU12009-AAU12024 represent recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 249 AA;

  Query Match             100.0%;  Score 1393;  DB 5;  Length 249;
  Best Local Similarity   100.0%;  Pred. No. 1e-122;
  Matches  249;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 2
AAU12017
ID   AAU12017 standard; protein; 249 AA.
XX
AC   AAU12017;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human beta-II tryptase active site mutant H44A #3.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20783.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 8; Page 105-106; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAU12009-AAU12024 represent recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 249 AA;

  Query Match             99.6%;  Score 1388;  DB 5;  Length 249;
  Best Local Similarity   99.6%;  Pred. No. 3e-122;
  Matches  248;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 3
AAY55011
ID   AAY55011 standard; protein; 249 AA.
XX
AC   AAY55011;
XX
DT   18-FEB-2000  (first entry)
XX
DE   Human beta-tryptase protein sequence.
XX
KW   Beta-tryptase; human; DNA expression construct; protein production;
KW   combinatorial library screening; X ray crystallography; antigen;
KW   antibody generation.
XX
OS   Homo sapiens.
XX
PN   WO9960139-A1.
XX
PD   25-NOV-1999.
XX
PF   29-OCT-1998;   98WO-US022994.
XX
PR   15-MAY-1998;   98US-00079970.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffitt MA,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2000-053300/04.
DR   N-PSDB; AAZ40175.
```

```
XX
PT   New DNA expression construct for production of enzymatically active
PT   recombinant human beta-tryptase.
XX
PS   Disclosure; Page 43-44; 50pp; English.
XX
CC   This sequence is the human beta-tryptase. The invention relates to a DNA
CC   expression construct comprising (5' to 3') a promoter linked to a signal
CC   sequence which is linked to a sequence encoding human beta-tryptase. The
CC   DNA construct is useful for transforming host cells to express, post
CC   translationally process and secrete enzymatically active human tryptase.
CC   The method is useful for the production of large amounts of tryptase with
CC   defined specifications. The transformant is useful for pharmacological
CC   studies, combinatorial library screens and X ray crystallographic
CC   studies. The tryptase produced allows for the development of tryptase
CC   agonists and/or antagonists, is useful as an antigen to generate
CC   antihuman tryptase antibodies in various animals, can be used in
CC   screening for compounds which act as tryptase inhibitors, antagonists,
CC   agonists etc. and to assay for the presence of tryptase in biological or
CC   other solutions. Tryptase inhibitors, antagonists, agonists etc. may be
CC   useful as therapeutics. The tryptase does not require any post-expression
CC   or post-purification modifications or manipulations to initiate tryptase
CC   activity and it has enzymatic activity which compares favourably with
CC   cadaveric tryptase. The availability of enzymatically active tryptase
CC   facilitates the large scale screening of combinatorial libraries for
CC   specific tryptase inhibitors as potential therapeutics and advances the
CC   understanding of the biological significance of tryptase in mast cell
CC   mediated diseases. The tryptase can be used to detect low levels of
CC   tryptase
XX
SQ   Sequence 249 AA;

  Query Match             99.6%;  Score 1387;  DB 3;  Length 249;
  Best Local Similarity   99.6%;  Pred. No. 3.7e-122;
  Matches  248;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 4
AAU12007
ID   AAU12007 standard; protein; 249 AA.
XX
AC   AAU12007;
XX
DT   09-APR-2002  (first entry)
XX
DE   Recombinant human beta-II tryptase.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutein; mutant.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
```

```
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20765.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 60; Page 77-78; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   The present sequence represents recombinant human beta-II tryptase
XX
SQ   Sequence 249 AA;

  Query Match             99.6%;  Score 1387;  DB 5;  Length 249;
  Best Local Similarity   99.6%;  Pred. No. 3.7e-122;
  Matches  248;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 5
AAU12011
ID   AAU12011 standard; protein; 249 AA.
XX
AC   AAU12011;
XX
DT   09-APR-2002  (first entry)
XX
```

```
DE   Human beta-II tryptase active site mutant S194A #1.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20777.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 8; Page 90-91; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAU12009-AAU12024 represent recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 249 AA;

  Query Match             99.4%;  Score 1384;  DB 5;  Length 249;
  Best Local Similarity   99.2%;  Pred. No. 7.1e-122;
  Matches  247;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249
```

```
RESULT 6
AAU12012
ID   AAU12012 standard; protein; 249 AA.
XX
AC   AAU12012;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human beta-II tryptase active site mutant S194A #2.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20778.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 8; Page 93; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAU12009-AAU12024 represent recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 249 AA;

  Query Match             99.4%;  Score 1384;  DB 5;  Length 249;
  Best Local Similarity   99.2%;  Pred. No. 7.1e-122;
  Matches  247;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
```

```
Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||:|||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 7
AAU12006
ID   AAU12006 standard; protein; 249 AA.
XX
AC   AAU12006;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human beta-I tryptase.
XX
KW   Human; proteolytic tryptase; protease; beta-I tryptase; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20763.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Disclosure; Page 74-75; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   The present sequence represents human beta-I tryptase
XX
SQ   Sequence 249 AA;

  Query Match             99.2%;  Score 1382;  DB 5;  Length 249;
  Best Local Similarity   99.2%;  Pred. No. 1.1e-121;
  Matches  247;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 8
AAU12020
ID   AAU12020 standard; protein; 249 AA.
XX
AC   AAU12020;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human beta-II tryptase active site mutant S194A #6.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20786.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 8; Page 114-115; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAU12009-AAU12024 represent recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 249 AA;
```

```
  Query Match            99.0%;  Score 1379;  DB 5;  Length 249;
  Best Local Similarity  98.8%;  Pred. No. 2.1e-121;
  Matches  246;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 9
AAU12010
ID   AAU12010 standard; protein; 249 AA.
XX
AC   AAU12010;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human beta-II tryptase active site mutant D91A #1.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20776.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 8; Page 87-88; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
```

```
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAU12009-AAU12024 represent recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 249 AA;

  Query Match            99.0%;  Score 1379;  DB 5;  Length 249;
  Best Local Similarity  99.2%;  Pred. No. 2.1e-121;
  Matches  247;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAAIALLELEEPVKVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 10
AAU12019
ID   AAU12019 standard; protein; 249 AA.
XX
AC   AAU12019;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human beta-II tryptase active site mutant S194A #5.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20785.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 8; Page 111-112; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
```

```
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAU12009-AAU12024 represent recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 249 AA;

  Query Match             99.0%;  Score 1379;  DB 5;  Length 249;
  Best Local Similarity   98.8%;  Pred. No. 2.1e-121;
  Matches  246;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              |||||||||||||||||:||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 11
AAU12013
ID   AAU12013 standard; protein; 245 AA.
XX
AC   AAU12013;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human beta-II tryptase active site mutant H44A #2.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20779.
```

```
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 40; Page 95-96; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAU12009-AAU12024 represent recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 245 AA;

  Query Match             98.6%;  Score 1374;  DB 5;  Length 245;
  Best Local Similarity   100.0%;  Pred. No. 6.1e-121;
  Matches  245;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKKP 249
              |||||
Db        241 VPKKP 245


RESULT 12
AAU12018
ID   AAU12018 standard; protein; 249 AA.
XX
AC   AAU12018;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human beta-II tryptase active site mutant D91A #3.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
```

```
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20784.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 8; Page 109-110; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAU12009-AAU12024 represent recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 249 AA;

  Query Match             98.6%;  Score 1374;  DB 5;  Length 249;
  Best Local Similarity   98.8%;  Pred. No. 6.2e-121;
  Matches  246;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60

Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              |||||||||||||||||||||||||||||||||| |||||||||| ||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAAIALLELEEPVNVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 13
AAU12021
ID   AAU12021 standard; protein; 245 AA.
XX
AC   AAU12021;
XX
DT   09-APR-2002  (first entry)
XX
DE   Human beta-II tryptase active site mutant H44A #4.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20787.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 40; Page 117-118; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAU12009-AAU12024 represent recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 245 AA;

  Query Match             98.3%;  Score 1369;  DB 5;  Length 245;
  Best Local Similarity   99.6%;  Pred. No. 1.8e-120;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKKP 249
              |||||
Db        241 VPKKP 245


RESULT 14
AAU12008
ID   AAU12008 standard; protein; 245 AA.
```

```
XX
AC   AAU12008;
XX
DT   09-APR-2002  (first entry)
XX
DE   Recombinant human mature beta-II tryptase.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutein; mutant.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   N-PSDB; AAS20766.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 61; Page 80; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   The present sequence represents recombinant human mature beta-II tryptase
XX
SQ   Sequence 245 AA;

  Query Match             98.2%;  Score 1368;  DB 5;  Length 245;
  Best Local Similarity   99.6%;  Pred. No. 2.2e-120;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
                ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240
```

```
Qy        245 VPKKP 249
              |||||
Db        241 VPKKP 245

RESULT 15
AAW64240
ID   AAW64240 standard; protein; 274 AA.
XX
AC   AAW64240;
XX
DT   24-NOV-1998  (first entry)
XX
DE   Human mast cell tryptase II/beta.
XX
KW   Mast cell tryptase II/beta; human; MCP-7; mast cell protease 7;
KW   blood clot; anticoagulant; myocardial infarction; reocclusion;
KW   thromboembolism; cerebral embolism; thrombosis; therapy.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1. .18
FT                   /label= Sig_peptide
FT   Peptide         19. .28
FT                   /label= Pro_peptide
FT   Protein         29. .274
FT                   /label= Mat_protein
XX
PN   WO9824886-A1.
XX
PD   11-JUN-1998.
XX
PF   25-NOV-1997;   97WO-US021620.
XX
PR   04-DEC-1996;   96US-0032354P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Stevens RL;
XX
DR   WPI; 1998-333308/29.
DR   N-PSDB; AAV44330.
XX
PT   New compositions containing tryptase-7, e.g. mouse mast cell protease-7 -
PT   are used to treat clot formation in e.g. myocardial infarction,
PT   reocclusion following angioplasty or pulmonary thrombo-embolism.
XX
PS   Disclosure; Page 66-67; 92pp; English.
XX
CC   This is the deduced amino acid sequence of human mast cell tryptase
CC   II/beta (see also AAV44330). The invention provides: compositions
CC   comprising an isolated tryptase-7 that may include chimeric proteins that
CC   contain (a) a human tryptase for all but the active site region and (b)
CC   the substrate-binding pocket of mouse tryptase-7 or its homologues (see
CC   AAW64233-39); a method for treating a blood clot by administering a
CC   nucleic acid molecule that codes for a tryptase-7, or an expression
CC   product, to decrease fibrinogen activity; a nucleic acid encoding a
CC   serine protease (SP); and a method of producing a mature SP by expressing
CC   the inactive zymogen in a host cell, and cleaving the enterokinase
CC   susceptibility domain. The tryptase-7 polypeptides can be used to treat
CC   disorders mediated by undesirable thrombus clot formation such as
CC   myocardial infarction and reocclusion following angioplasty of blood
CC   clots associated with pulmonary thromboembolism, deep vein thrombosis,
CC   cerebral embolism, renal vein and peripheral arterial thrombosis. They
CC   are also useful for all surgical procedures that require decreased blood
CC   clots
XX
SQ   Sequence 274 AA;

  Query Match             98.2%;  Score 1368;  DB 2;  Length 274;
  Best Local Similarity   99.6%;  Pred. No. 2.6e-120;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         30 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 89

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 149

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 209

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 269

Qy        245 VPKKP 249
              |||||
Db        270 VPKKP 274

Search completed: July 22, 2004, 15:22:25
Job time : 100 secs
```

```
                       GenCore version 5.1.6
              Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          July 22, 2004, 15:20:41 ; Search time 58 Seconds
                                           (without alignments)
                                           221.636 Million cell updates/sec

Title:           US-09-598-982-21
Perfect score:   1393
Sequence:        1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters:      389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :        Issued_Patents_AA:*
                 1:  /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
                 2:  /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
                 3:  /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
                 4:  /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
                 5:  /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
                 6:  /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB ID                           Description
-------------------------------------------------------------------------------
    1    1387    99.6    249  3  US-09-079-970A-5             Sequence 5, Appli
    2    1368    98.2    245  3  US-09-079-970A-6             Sequence 6, Appli
    3    1368    98.2    245  4  US-09-601-318-1              Sequence 1, Appli
    4    1368    98.2    274  2  US-09-016-366A-21            Sequence 21, Appl
    5    1368    98.2    274  2  US-08-978-404B-16            Sequence 16, Appl
    6    1363    97.8    273  2  US-09-016-366A-19            Sequence 19, Appl
    7    1363    97.8    273  2  US-08-978-404B-14            Sequence 14, Appl
    8    1361    97.7    244  4  US-09-601-318-4              Sequence 4, Appli
    9    1361    97.7    244  4  US-09-601-318-5              Sequence 5, Appli
   10    1361    97.7    244  4  US-09-601-318-6              Sequence 6, Appli
   11    1361    97.7    244  4  US-09-601-318-7              Sequence 7, Appli
   12    1344    96.5    267  2  US-09-016-366A-23            Sequence 23, Appl
   13    1344    96.5    267  2  US-08-978-404B-18            Sequence 18, Appl
   14    1274    91.5    245  3  US-08-944-483-69             Sequence 69, Appl
   15    1274    91.5    275  2  US-09-016-366A-17            Sequence 17, Appl
   16    1274    91.5    275  2  US-08-978-404B-12            Sequence 12, Appl
   17    1091    78.3    276  2  US-09-016-366A-15            Sequence 15, Appl
   18    1091    78.3    276  2  US-08-978-404B-21            Sequence 21, Appl
   19    1075    77.2    270  2  US-08-978-404B-8             Sequence 8, Appli
   20    1054    75.7    273  2  US-08-978-404B-3             Sequence 3, Appli
   21    1051    75.4    274  2  US-08-978-404B-5             Sequence 5, Appli
   22    1037    74.4    273  2  US-08-978-404B-6             Sequence 6, Appli
   23     702    50.4    190  2  US-08-845-998-4              Sequence 4, Appli
   24     702    50.4    190  3  US-09-206-537-4              Sequence 4, Appli
   25     702    50.4    190  3  US-09-430-854-4              Sequence 4, Appli
   26     698    50.1    190  2  US-08-845-998-6              Sequence 6, Appli
   27     698    50.1    190  3  US-09-206-537-6              Sequence 6, Appli
```

```
   28     698    50.1    190  3  US-09-430-854-6              Sequence 6, Appli
   29   684.5    49.1    269  2  US-08-978-404B-10            Sequence 10, Appl
   30   563.5    40.5    290  4  US-09-386-653A-7             Sequence 7, Appli
   31   554.5    39.8    315  4  US-09-386-653A-9             Sequence 9, Appli
   32   536.5    38.5    284  4  US-09-387-375-7              Sequence 7, Appli
   33   527.5    37.9    316  4  US-09-387-375-9              Sequence 9, Appli
   34     507    36.4    319  4  US-09-386-642-12             Sequence 12, Appl
   35     507    36.4    328  4  US-09-386-642-11             Sequence 11, Appl
   36     506    36.3    299  3  US-08-944-483-66             Sequence 66, Appl
   37   501.5    36.0    317  4  US-09-386-629-7              Sequence 7, Appli
   38   501.5    36.0    317  4  US-09-907-794A-263           Sequence 263, App
   39   501.5    36.0    317  4  US-09-905-125A-263           Sequence 263, App
   40   501.5    36.0    317  4  US-09-902-775A-263           Sequence 263, App
   41   499.5    35.9    312  4  US-09-023-942A-4             Sequence 4, Appli
   42   493.5    35.4    314  3  US-09-008-271A-3             Sequence 3, Appli
   43   493.5    35.4    314  4  US-09-907-794A-257           Sequence 257, App
   44   493.5    35.4    314  4  US-09-905-125A-257           Sequence 257, App
   45   493.5    35.4    314  4  US-09-902-775A-257           Sequence 257, App

                                    ALIGNMENTS


RESULT 1
US-09-079-970A-5
; Sequence 5, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  249 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-079-970A-5

  Query Match             99.6%;  Score 1387;  DB 3;  Length 249;
  Best Local Similarity   99.6%;  Pred. No. 5.5e-141;
  Matches  248;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAAL 60
              ||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAAL 60
```

```
Qy         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 RVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240

Qy        241 IHHYVPKKP 249
              |||||||||
Db        241 IHHYVPKKP 249


RESULT 2
US-09-079-970A-6
; Sequence 6, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  245 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-079-970A-6

  Query Match              98.2%;  Score 1368;  DB 3;  Length 245;
  Best Local Similarity    99.6%;  Pred. No. 5.9e-139;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKKP 249
              |||||
Db        241 VPKKP 245


RESULT 3
US-09-601-318-1
; Sequence 1, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 245
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-1

  Query Match              98.2%;  Score 1368;  DB 4;  Length 245;
  Best Local Similarity    99.6%;  Pred. No. 5.9e-139;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKKP 249
              |||||
Db        241 VPKKP 245
```

```
RESULT 4
US-09-016-366A-21
; Sequence 21, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  21:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  274 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-016-366A-21

  Query Match             98.2%;  Score 1368;  DB 2;  Length 274;
  Best Local Similarity   99.6%;  Pred. No. 6.9e-139;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         30 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 89

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 149

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 209

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 269

Qy        245 VPKKP 249
              |||||
Db        270 VPKKP 274

RESULT 5
US-08-978-404B-16
; Sequence 16, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  274 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5968782e
US-08-978-404B-16

  Query Match             98.2%;  Score 1368;  DB 2;  Length 274;
  Best Local Similarity   99.6%;  Pred. No. 6.9e-139;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         30 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 89

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         90 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 149

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        150 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 209

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        210 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 269

Qy        245 VPKKP 249
              |||||
Db        270 VPKKP 274

RESULT 6
US-09-016-366A-19
; Sequence 19, Application US/09016366A
; Patent No. 5955431
```

```
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  273 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-016-366A-19

  Query Match            97.8%;  Score 1363;  DB 2;  Length 273;
  Best Local Similarity  99.2%;  Pred. No. 2.4e-138;
  Matches  243;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         29 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 88

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         89 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFP 148

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 208

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 268

Qy        245 VPKKP 249
              |||||
Db        269 VPKKP 273


RESULT 7
US-08-978-404B-14
; Sequence 14, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
```

```
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  273 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5968782e
US-08-978-404B-14

  Query Match            97.8%;  Score 1363;  DB 2;  Length 273;
  Best Local Similarity  99.2%;  Pred. No. 2.4e-138;
  Matches  243;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         29 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 88

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         89 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFP 148

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 208

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 268

Qy        245 VPKKP 249
              |||||
Db        269 VPKKP 273


RESULT 8
US-09-601-318-4
; Sequence 4, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
```

```
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-4

  Query Match             97.7%;  Score 1361;  DB 4;  Length 244;
  Best Local Similarity   99.6%;  Pred. No. 3.3e-138;
  Matches  243;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKK 248
              ||||
Db        241 VPKK 244

RESULT 9
US-09-601-318-5
; Sequence 5, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-5

  Query Match             97.7%;  Score 1361;  DB 4;  Length 244;
  Best Local Similarity   99.6%;  Pred. No. 3.3e-138;
  Matches  243;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKK 248
              ||||
Db        241 VPKK 244

RESULT 10
US-09-601-318-6
; Sequence 6, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
```

```
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-601-318-6

  Query Match             97.7%;  Score 1361;  DB 4;  Length 244;
  Best Local Similarity   99.6%;  Pred. No. 3.3e-138;
  Matches  243;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKK 248
              ||||
Db        241 VPKK 244

RESULT 11
US-09-601-318-7
; Sequence 7, Application US/09601318
; Patent No. 6613769
; GENERAL INFORMATION:
;  APPLICANT: Max-Planck-Gesellschaft z. Frd. d. Wissensch. e.V
;  APPLICANT:  Byk Gulden Lomberg Chemische Fabrik GmbH
;  APPLICANT:  Bode, Wolfram
;  APPLICANT:  Moroder, Luis
;  APPLICANT:  Pereira, Pedro Jose Barbosa
;  APPLICANT:  Bergner, Andreas
;  APPLICANT:  Huber, Robert
;  APPLICANT:  Sommerhoff, Christian
;  APPLICANT:  Schaschke, No. 6613769bert
;  APPLICANT:  Br, Thomas
;  APPLICANT:  Martin, Thomas
;  APPLICANT:  Stadlwieser, Josef
;  APPLICANT:  Ulrich, Wolf-Rdiger
;  APPLICANT:  Dominik, Andreas
;  APPLICANT:  Thibaut, Ulrich
;  APPLICANT:  Bundschuh, Daniela
;  APPLICANT:  Beume, Rolf
;  APPLICANT:  Goebel, Karl-Josef
;  TITLE OF INVENTION: Tryptase-Inhibitoren
;  FILE REFERENCE: 17674P WO-1
;  CURRENT APPLICATION NUMBER: US/09/601,318
;  CURRENT FILING DATE:  2001-01-22
;  PRIOR APPLICATION NUMBER: DE19804761.4
;  PRIOR FILING DATE: 1998-02-06
;  PRIOR APPLICATION NUMBER: DE19851300.3
;  PRIOR FILING DATE: 1998-11-06
;  NUMBER OF SEQ ID NOS: 7
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 244
;   TYPE: PRT
;   ORGANISM: Homo sapiens
```

```
US-09-601-318-7

  Query Match             97.7%;  Score 1361;  DB 4;  Length 244;
  Best Local Similarity   99.6%;  Pred. No. 3.3e-138;
  Matches  243;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          1 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKK 248
              ||||
Db        241 VPKK 244

RESULT 12
US-09-016-366A-23
; Sequence 23, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  23:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  267 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-016-366A-23

  Query Match             96.5%;  Score 1344;  DB 2;  Length 267;
```

```
  Best Local Similarity   98.4%;  Pred. No. 2.5e-136;
  Matches  241;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| ||||||||||||||||
Db         23 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 82

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 142

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 202

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 262

Qy        245 VPKKP 249
              |||||
Db        263 VPKKP 267

RESULT 13
US-08-978-404B-18
; Sequence 18, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  267 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 5968782e
US-08-978-404B-18

  Query Match             96.5%;  Score 1344;  DB 2;  Length 267;
  Best Local Similarity   98.4%;  Pred. No. 2.5e-136;
  Matches  241;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| ||||||||||||||||
Db         23 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 82

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         83 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 142

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        143 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 202

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        203 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 262

Qy        245 VPKKP 249
              |||||
Db        263 VPKKP 267

RESULT 14
US-08-944-483-69
; Sequence 69, Application US/08944483
; Patent No. 6232456
;  GENERAL INFORMATION:
;    APPLICANT:  COHEN, MAURICE
;    APPLICANT:  COLPITTS, TRACEY L.
;    APPLICANT:  FRIEDMAN, PAULA N.
;    APPLICANT:  GRANADOS, EDWARD N.
;    APPLICANT:  KLASS, MICHAEL R.
;    APPLICANT:  RUSSELL, JOHN C.
;    APPLICANT:  STEWART, KENT D.
;    APPLICANT:  STROUPE, STEVEN D.
;    TITLE OF INVENTION:  NOVEL SERINE PROTEASE REAGENTS
;    TITLE OF INVENTION:  AND METHODS USEFUL FOR DETECTING AND TREATING DISEASES
;    TITLE OF INVENTION:  OF THE PROSTATE
;    NUMBER OF SEQUENCES:  76
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Abbott Laboratories
;      STREET:  100 Abbott Park Road
;      CITY:  Abbott Park
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60064-3500
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/944,483
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Becker, Cheryl L.
;      REGISTRATION NUMBER:  35,441
;      REFERENCE/DOCKET NUMBER:  6183.US.01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  847/935-1729
;      TELEFAX:  847/938-2623
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  69:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  245 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  No. 6232456e
US-08-944-483-69
```

```
  Query Match             91.5%;  Score 1274;  DB 3;  Length 245;
  Best Local Similarity   92.2%;  Pred. No. 7.5e-129;
  Matches  226;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| |:|||||||| |||||
Db          1 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 60

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db         61 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 120

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db        121 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 180

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||::||||:||||||||||||||||||||||| |||||||||||||||||||||||||
Db        181 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 240

Qy        245 VPKKP 249
              |||||
Db        241 VPKKP 245

RESULT 15
US-09-016-366A-17
; Sequence 17, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  275 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-016-366A-17

  Query Match             91.5%;  Score 1274;  DB 2;  Length 275;
  Best Local Similarity   92.2%;  Pred. No. 8.8e-129;
  Matches  226;  Conservative    6;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| |:|||||||| |||||
Db         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db        151 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||::||||:||||||||||||||||||||||| |||||||||||||||||||||||||
Db        211 CAGNSQRDSCKGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275

Search completed: July 22, 2004, 15:27:27
Job time : 59 secs
```

```
                 GenCore version 5.1.6
        Copyright (c) 1993 - 2004  Compugen Ltd.


OM protein - protein search, using sw model

Run on:         July 22, 2004, 15:25:12 ; Search time 940 Seconds
                                           (without alignments)
                                           82.953 Million cell updates/sec

Title:          US-09-598-982-21
Perfect score:  1393
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched:       1288442 seqs, 313154207 residues

Total number of hits satisfying chosen parameters:      1288442

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_AA:*
               1:  /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
               2:  /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
               3:  /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
               4:  /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
               5:  /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
               6:  /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
               7:  /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
               8:  /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
               9:  /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
               10:  /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
               11:  /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
               12:  /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
               13:  /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
               14:  /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
               15:  /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
               16:  /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
               17:  /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
               18:  /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                %
Result          Query
  No.   Score   Match Length DB  ID                          Description
 -----------------------------------------------------------------------------
    1    1368    98.2    254  12  US-10-150-813-74            Sequence 74, Appl
    2    1368    98.2    254  12  US-10-139-854-74            Sequence 74, Appl
    3    1368    98.2    254  14  US-10-131-409-74            Sequence 74, Appl
    4    1368    98.2    254  15  US-10-150-811-74            Sequence 74, Appl
    5    1368    98.2    256  12  US-10-150-813-92            Sequence 92, Appl
    6    1368    98.2    256  12  US-10-139-854-92            Sequence 92, Appl
    7    1368    98.2    256  14  US-10-131-409-92            Sequence 92, Appl
    8    1368    98.2    256  15  US-10-150-811-92            Sequence 92, Appl
    9    1368    98.2    264  12  US-10-150-813-104           Sequence 104, App
   10    1368    98.2    264  12  US-10-139-854-104           Sequence 104, App
   11    1368    98.2    264  14  US-10-131-409-104           Sequence 104, App
   12    1368    98.2    264  15  US-10-150-811-104           Sequence 104, App
   13    1368    98.2    275  10  US-09-813-432-44            Sequence 44, Appl
   14    1368    98.2    275  12  US-10-246-583-44            Sequence 44, Appl
   15    1368    98.2    275  15  US-10-352-684A-48           Sequence 48, Appl
   16    1368    98.2    275  15  US-10-174-364-44            Sequence 44, Appl
   17    1368    98.2    275  16  US-10-287-226-94            Sequence 94, Appl
   18    1368    98.2    275  16  US-10-689-832-44            Sequence 44, Appl
   19    1363    97.8    275  16  US-10-287-226-92            Sequence 92, Appl
   20  1347.5    96.7    272  14  US-10-117-323-40            Sequence 40, Appl
   21    1329    95.4    691  16  US-10-275-505-13            Sequence 13, Appl
   22    1323    95.0    275  16  US-10-287-226-90            Sequence 90, Appl
   23    1317    94.5    266  12  US-10-150-813-84            Sequence 84, Appl
   24    1317    94.5    266  12  US-10-139-854-84            Sequence 84, Appl
   25    1317    94.5    266  14  US-10-131-409-84            Sequence 84, Appl
   26    1317    94.5    266  15  US-10-150-811-84            Sequence 84, Appl
   27    1312    94.2    279  16  US-10-287-226-88            Sequence 88, Appl
   28    1274    91.5    275  14  US-10-117-323-36            Sequence 36, Appl
   29    1274    91.5    275  15  US-10-352-684A-46           Sequence 46, Appl
   30    1270    91.2    231  10  US-09-898-837A-52           Sequence 52, Appl
   31    1104    79.3    199  12  US-10-150-813-98            Sequence 98, Appl
   32    1104    79.3    199  12  US-10-139-854-98            Sequence 98, Appl
   33    1104    79.3    199  14  US-10-131-409-98            Sequence 98, Appl
   34    1104    79.3    199  15  US-10-150-811-98            Sequence 98, Appl
   35    1091    78.3    252  10  US-09-813-432-42            Sequence 42, Appl
   36    1091    78.3    252  12  US-10-246-583-42            Sequence 42, Appl
   37    1091    78.3    252  15  US-10-174-364-42            Sequence 42, Appl
   38    1091    78.3    252  16  US-10-689-832-42            Sequence 42, Appl
   39    1091    78.3    276  14  US-10-117-323-35            Sequence 35, Appl
   40    1077    77.3    276  14  US-10-311-955-3             Sequence 3, Appli
   41    1064    76.4    247  12  US-10-150-813-86            Sequence 86, Appl
   42    1064    76.4    247  12  US-10-139-854-86            Sequence 86, Appl
   43    1064    76.4    247  14  US-10-131-409-86            Sequence 86, Appl
   44    1064    76.4    247  15  US-10-150-811-86            Sequence 86, Appl
   45    1054    75.7    310  16  US-10-275-505-8             Sequence 8, Appli


                                  ALIGNMENTS


RESULT 1
US-10-150-813-74
; Sequence 74, Application US/10150813
; Publication No. US20030224367A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP1CON2
;  CURRENT APPLICATION NUMBER: US/10/150,813
;  CURRENT FILING DATE:  2002-12-03
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
;   LENGTH: 254
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-813-74
```

```
  Query Match             98.2%;  Score 1368;  DB 12;  Length 254;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         10 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 69

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 129

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 189

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 249

Qy        245 VPKKP 249
              |||||
Db        250 VPKKP 254


RESULT 2
US-10-139-854-74
; Sequence 74, Application US/10139854
; Publication No. US20030202971A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder, Kumud
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CON2
;  CURRENT APPLICATION NUMBER: US/10/139,854
;  CURRENT FILING DATE:  2002-12-02
;  PRIOR APPLICATION NUMBER: 09/783,429
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 126
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
;   LENGTH: 254
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-139-854-74

  Query Match             98.2%;  Score 1368;  DB 12;  Length 254;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         10 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 69

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 129

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 189

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 249

Qy        245 VPKKP 249
              |||||
Db        250 VPKKP 254


RESULT 3
US-10-131-409-74
; Sequence 74, Application US/10131409
; Publication No. US20030199465A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20030199465A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP1CON1
;  CURRENT APPLICATION NUMBER: US/10/131,409
;  CURRENT FILING DATE:  2002-10-24
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
;   LENGTH: 254
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-131-409-74

  Query Match             98.2%;  Score 1368;  DB 14;  Length 254;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         10 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 69

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 129

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        130 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 189

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         190 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 249

Qy         245 VPKKP 249
               |||||
Db         250 VPKKP 254


RESULT 4
US-10-150-811-74
; Sequence 74, Application US/10150811
; Publication No. US20040010120A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20040010120A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 15966-675CIP2CON1
;  CURRENT APPLICATION NUMBER: US/10/150,811
;  CURRENT FILING DATE:  2002-05-17
;  PRIOR APPLICATION NUMBER: 09/970,607
;  PRIOR FILING DATE: 2001-10-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
;   LENGTH: 254
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-811-74

  Query Match             98.2%;  Score 1368;  DB 15;  Length 254;
  Best Local Similarity   99.6%;  Pred. No. 1.5e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy           5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
               ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          10 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 69

Qy          65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          70 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 129

Qy         125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         130 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 189

Qy         185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         190 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 249

Qy         245 VPKKP 249
               |||||
Db         250 VPKKP 254


RESULT 5
US-10-150-813-92
; Sequence 92, Application US/10150813
; Publication No. US20030224367A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP1CON2
;  CURRENT APPLICATION NUMBER: US/10/150,813
;  CURRENT FILING DATE:  2002-12-03
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-813-92

  Query Match             98.2%;  Score 1368;  DB 12;  Length 256;
  Best Local Similarity   99.6%;  Pred. No. 1.6e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy           5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
               ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db          12 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 71

Qy          65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          72 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 131

Qy         125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         132 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 191

Qy         185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         192 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 251

Qy         245 VPKKP 249
               |||||
Db         252 VPKKP 256


RESULT 6
US-10-139-854-92
; Sequence 92, Application US/10139854
; Publication No. US20030202971A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder, Kumud
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CON2
;  CURRENT APPLICATION NUMBER: US/10/139,854
;  CURRENT FILING DATE:  2002-12-02
```

```
;  PRIOR APPLICATION NUMBER: 09/783,429
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 126
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-139-854-92

  Query Match             98.2%;  Score 1368;  DB 12;  Length 256;
  Best Local Similarity   99.6%;  Pred. No. 1.6e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         12 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 71

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 131

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 191

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 251

Qy        245 VPKKP 249
              |||||
Db        252 VPKKP 256


RESULT 7
US-10-131-409-92
; Sequence 92, Application US/10131409
; Publication No. US20030199465A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20030199465A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 15966-675CIP1CON1
;  CURRENT APPLICATION NUMBER: US/10/131,409
;  CURRENT FILING DATE:  2002-10-24
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
```

```
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-131-409-92

  Query Match             98.2%;  Score 1368;  DB 14;  Length 256;
  Best Local Similarity   99.6%;  Pred. No. 1.6e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         12 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 71

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 131

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 191

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 251

Qy        245 VPKKP 249
              |||||
Db        252 VPKKP 256


RESULT 8
US-10-150-811-92
; Sequence 92, Application US/10150811
; Publication No. US20040010120A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20040010120A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP2CON1
;  CURRENT APPLICATION NUMBER: US/10/150,811
;  CURRENT FILING DATE:  2002-05-17
;  PRIOR APPLICATION NUMBER: 09/970,607
;  PRIOR FILING DATE: 2001-10-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
```

```
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 92
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-811-92

  Query Match             98.2%;  Score 1368;  DB 15;  Length 256;
  Best Local Similarity   99.6%;  Pred. No. 1.6e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         12 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 71

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         72 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 131

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        132 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 191

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        192 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 251

Qy        245 VPKKP 249
              |||||
Db        252 VPKKP 256

RESULT 9
US-10-150-813-104
; Sequence 104, Application US/10150813
; Publication No. US20030224367A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP1CON2
;  CURRENT APPLICATION NUMBER: US/10/150,813
;  CURRENT FILING DATE:  2002-12-03
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
;   LENGTH: 264
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-813-104

  Query Match             98.2%;  Score 1368;  DB 12;  Length 264;
  Best Local Similarity   99.6%;  Pred. No. 1.6e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         20 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 79

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 139

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 199

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 259

Qy        245 VPKKP 249
              |||||
Db        260 VPKKP 264

RESULT 10
US-10-139-854-104
; Sequence 104, Application US/10139854
; Publication No. US20030202971A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder, Kumud
;  TITLE OF INVENTION: Novel Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CON2
;  CURRENT APPLICATION NUMBER: US/10/139,854
;  CURRENT FILING DATE:  2002-12-02
;  PRIOR APPLICATION NUMBER: 09/783,429
;  PRIOR FILING DATE: 2001-02-14
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 126
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
;   LENGTH: 264
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-139-854-104

  Query Match             98.2%;  Score 1368;  DB 12;  Length 264;
  Best Local Similarity   99.6%;  Pred. No. 1.6e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         20 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 79
```

```
Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 139

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 199

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 259

Qy        245 VPKKP 249
              |||||
Db        260 VPKKP 264


RESULT 11
US-10-131-409-104
; Sequence 104, Application US/10131409
; Publication No. US20030199465A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20030199465A1el Polypeptides and Nucleic Acids Encoding Sam
;  FILE REFERENCE: 15966-675CIP1CON1
;  CURRENT APPLICATION NUMBER: US/10/131,409
;  CURRENT FILING DATE:  2002-10-24
;  PRIOR APPLICATION NUMBER: 09/898,954
;  PRIOR FILING DATE: 2001-07-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 135
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
;   LENGTH: 264
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-131-409-104

  Query Match             98.2%;  Score 1368;  DB 14;  Length 264;
  Best Local Similarity   99.6%;  Pred. No. 1.6e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         20 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 79

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 139

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 199
```

```
Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 259

Qy        245 VPKKP 249
              |||||
Db        260 VPKKP 264

RESULT 12
US-10-150-811-104
; Sequence 104, Application US/10150811
; Publication No. US20040010120A1
; GENERAL INFORMATION:
;  APPLICANT: Malyankar et al.
;  TITLE OF INVENTION: No. US20040010120A1el Polypeptides and Nucleic Acids Encoding Same
;  FILE REFERENCE: 15966-675CIP2CON1
;  CURRENT APPLICATION NUMBER: US/10/150,811
;  CURRENT FILING DATE:  2002-05-17
;  PRIOR APPLICATION NUMBER: 09/970,607
;  PRIOR FILING DATE: 2001-10-03
;  PRIOR APPLICATION NUMBER: 60/182,733
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/182,724
;  PRIOR FILING DATE: 2000-02-15
;  PRIOR APPLICATION NUMBER: 60/183,896
;  PRIOR FILING DATE: 2000-02-22
;  PRIOR APPLICATION NUMBER: 60/184,497
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/224,157
;  PRIOR FILING DATE: 2000-08-10
;  PRIOR APPLICATION NUMBER: 60/184,482
;  PRIOR FILING DATE: 2000-02-23
;  PRIOR APPLICATION NUMBER: 60/184,744
;  PRIOR FILING DATE: 2000-02-24
;  PRIOR APPLICATION NUMBER: 60/197,083
;  PRIOR FILING DATE: 2000-04-13
;  PRIOR APPLICATION NUMBER: 60/233,405
;  PRIOR FILING DATE: 2000-09-18
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 138
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
;   LENGTH: 264
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-150-811-104

  Query Match             98.2%;  Score 1368;  DB 15;  Length 264;
  Best Local Similarity   99.6%;  Pred. No. 1.6e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         20 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 79

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         80 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 139

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        140 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 199

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        200 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 259

Qy        245 VPKKP 249
              |||||
Db        260 VPKKP 264
```

```
RESULT 13
US-09-813-432-44
; Sequence 44, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier Jr., Raymond J
;  APPLICANT:  Majmuder, Kamud
;  APPLICANT:  Spaderna, Steven K
;  APPLICANT:  Smithson, Glenda
;  APPLICANT:  Mezes, Peter S
;  APPLICANT:  Vernet, Corine A. M.
;  TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
;  FILE REFERENCE: 15966-729
;  CURRENT APPLICATION NUMBER: US/09/813,432
;  CURRENT FILING DATE:  2001-03-20
;  PRIOR APPLICATION NUMBER: 60/190,835
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,768
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,972
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,199
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,947
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,665
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,657
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,984
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,664
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,836
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: 60/193,843
;  PRIOR FILING DATE: 2000-03-31
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-813-432-44

  Query Match             98.2%;  Score 1368;  DB 10;  Length 275;
  Best Local Similarity   99.6%;  Pred. No. 1.7e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275


RESULT 14
US-10-246-583-44
; Sequence 44, Application US/10246583
```

```
; Publication No. US20040058862A1
; GENERAL INFORMATION:
;  APPLICANT: Majumder
;  TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 15966-729CIP2CON1
;  CURRENT APPLICATION NUMBER: US/10/246,583
;  CURRENT FILING DATE:  2002-12-06
;  PRIOR APPLICATION NUMBER: 10/174,364
;  PRIOR FILING DATE: 2002-06-17
;  PRIOR APPLICATION NUMBER: 60/190,835
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,768
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,972
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,199
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,947
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,665
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,657
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,984
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,664
;  PRIOR FILING DATE: 2000-03-28
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 128
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 44
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-246-583-44

  Query Match             98.2%;  Score 1368;  DB 12;  Length 275;
  Best Local Similarity   99.6%;  Pred. No. 1.7e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275


RESULT 15
US-10-352-684A-48
; Sequence 48, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
```

```
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 48
;   LENGTH: 275
;   TYPE: PRT
;   ORGANISM: Homo Sapiens
US-10-352-684A-48

  Query Match             98.2%;  Score 1368;  DB 15;  Length 275;
  Best Local Similarity   99.6%;  Pred. No. 1.7e-124;
  Matches  244;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||||||||||||||||||||||||||||||||||||||||| |||||||||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275

Search completed: July 22, 2004, 15:54:31
Job time : 941 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 22, 2004, 15:27:32 ; Search time 4800 Seconds
                                          (without alignments)
                                          2248.420 Million cell updates/sec

Title:          US-09-598-982-21
Perfect score:  1393
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters:      6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_20072004_064409_6487/app_query.fasta_1.391
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982_@CGN_1_1_7406_@runat_20072004_064409_6487 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                   SUMMARIES
                      %
Result              Query
  No.     Score     Match  Length  DB  ID                 Description
-----------------------------------------------------------------------------
   1      1393      100.0     771   6  AX347853           AX347853 Sequence
   2      1388       99.6     771   6  AX347869           AX347869 Sequence
   3      1387       99.6     771   6  AR165112           AR165112 Sequence
   4      1387       99.6     771   6  AX347841           AX347841 Sequence
   5      1387       99.6     771   6  BD211490           BD211490 Enzymatic
   6      1384       99.4     771   6  AX347857           AX347857 Sequence
   7      1384       99.4     771   6  AX347859           AX347859 Sequence
   8      1382       99.2     771   6  AX347838           AX347838 Sequence
   9      1379       99.0     771   6  AX347855           AX347855 Sequence
  10      1379       99.0     771   6  AX347873           AX347873 Sequence
  11      1379       99.0     771   6  AX347875           AX347875 Sequence
  12      1374       98.6     735   6  AX347861           AX347861 Sequence
  13      1374       98.6     771   6  AX347871           AX347871 Sequence
  14      1369       98.3     735   6  AX347877           AX347877 Sequence
  15      1368       98.2     735   6  AR165109           AR165109 Sequence
  16      1368       98.2     735   6  AX347843           AX347843 Sequence
  17      1368       98.2     735   6  BD211487           BD211487 Enzymatic
  18      1368       98.2    1128   6  AR080461           AR080461 Sequence
  19      1368       98.2    1128   9  HUMTRY2A           M33492 Human trypt
  20      1368       98.2    1143   9  HUMBTRYP           M37488 Human beta-
  21      1368       98.2    1145   9  S55551             S55551 beta-trypta
  22      1368       98.2    1194   9  BC029356           BC029356 Homo sapi
  23      1368       98.2    1235   9  BC038416           BC038416 Homo sapi
  24      1365       98.0     735   6  AX347865           AX347865 Sequence
  25      1365       98.0     735   6  AX347867           AX347867 Sequence
  26      1363       97.8     735   6  AX347834           AX347834 Sequence
  27      1363       97.8    1137   6  AR080460           AR080460 Sequence
  28      1363       97.8    1137   9  HUMTRY1A           M33491 Human trypt
  29      1360       97.6     735   6  AX347863           AX347863 Sequence
  30      1360       97.6     735   6  AX347881           AX347881 Sequence
  31      1360       97.6     735   6  AX347883           AX347883 Sequence
  32      1355       97.3     735   6  AX347879           AX347879 Sequence
  33      1344       96.5    1081   6  AR080462           AR080462 Sequence
  34      1344       96.5    1081   6  AX329646           AX329646 Sequence
  35      1344       96.5    1081   6  AX330205           AX330205 Sequence
  36      1344       96.5    1081   6  AX334644           AX334644 Sequence
  37      1344       96.5    1081   6  AX375047           AX375047 Sequence
  38      1344       96.5    1081   9  HUMTRY3A           M33493 Human trypt
  39      1341       96.3    1221   6  AX014344           AX014344 Sequence
  40      1341       96.3    1221   6  BD222223           BD222223 Human nuc
  41      1329       95.4    2662   6  AX338508           AX338508 Sequence
  42      1286       92.3     884   9  AF206665           AF206665 Homo sapi
  43      1286       92.3    1206   9  BC028059           BC028059 Homo sapi
  44      1286       92.3    1207   9  BC051852           BC051852 Homo sapi
  45      1274       91.5    1154   6  AR080459           AR080459 Sequence


                                   ALIGNMENTS

RESULT 1
```

```
AX347853
LOCUS       AX347853                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 20 from Patent WO0198470.
ACCESSION   AX347853
VERSION     AX347853.1  GI:18495660
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 20 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <1..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22420.1"
                     /db_xref="GI:18495661"
                     /db_xref="REMTREMBL:CAD22420"
                     /translation="GPLEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQW
                     VLTAAACVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLE
                     LEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIME
                     NHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVV
                     SWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              1.77e-122      Length:         771
Score:                  1393.00        Matches:        249
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     6              Gaps:           0

US-09-598-982-21 (1-249) x AX347853 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 2
AX347869
LOCUS       AX347869                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 36 from Patent WO0198470.
ACCESSION   AX347869
VERSION     AX347869.1  GI:18495676
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 36 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22428.1"
                     /db_xref="GI:18495677"
                     /db_xref="REMTREMBL:CAD22428"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAACVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              5.25e-122      Length:         771
Score:                  1388.00        Matches:        248
Percent Similarity:     99.60%         Conservative:   0
Best Local Similarity:  99.60%         Mismatches:     1
Query Match:            99.64%         Indels:         0
DB:                     6              Gaps:           0

US-09-598-982-21 (1-249) x AX347869 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
```

```
                   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy             41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            127 CAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy             61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy             81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy            101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
                  |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db            307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy            121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy            141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy            161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy            181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy            201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy            221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy            241 IleHisHisTyrValProLysLysPro 249
                  |||||||||||||||||||||||||||
Db            727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 3
AR165112
LOCUS       AR165112                 771 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 4 from patent US 6274366.
ACCESSION   AR165112
VERSION     AR165112.1  GI:16238527
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 771)
  AUTHORS   Maffitt,M.A., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Enzymatically-active recombinant human .beta.-tryptase and method
            of making same
  JOURNAL   Patent: US 6274366-A 4 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              6.53e-122      Length:         771
Score:                  1387.00        Matches:        248
Percent Similarity:     99.60%         Conservative:   0
Best Local Similarity:  99.60%         Mismatches:     1
```

```
Query Match:            99.57%         Indels:         0
DB:                     6              Gaps:           0

US-09-598-982-21 (1-249) x AR165112 (1-771)

Qy              1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db              7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy             21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy             41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
                  |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db            127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy             61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy             81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy            101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy            121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy            141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy            161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy            181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy            201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy            221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
                  ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy            241 IleHisHisTyrValProLysLysPro 249
                  |||||||||||||||||||||||||||
Db            727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 4
AX347841
LOCUS       AX347841                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 8 from Patent WO0198470.
ACCESSION   AX347841
VERSION     AX347841.1  GI:18495648
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 8 27-DEC-2001;
```

```
                PROMEGA CORPORATION (US)
FEATURES                 Location/Qualifiers
     source              1. .771
                         /organism="Homo sapiens"
                         /mol_type="unassigned DNA"
                         /db_xref="taxon:9606"
     CDS                 <7. .756
                         /note="unnamed protein product"
                         /codon_start=1
                         /protein_id="CAD22418.1"
                         /db_xref="GI:18495649"
                         /db_xref="REMTREMBL:CAD22418"
                         /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                         TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                         EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                         ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSW
                         GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              6.53e-122       Length:         771
Score:                  1387.00         Matches:        248
Percent Similarity:     99.60%          Conservative:   0
Best Local Similarity:  99.60%          Mismatches:     1
Query Match:            99.57%          Indels:         0
DB:                     6               Gaps:           0

US-09-598-982-21 (1-249) x AX347841 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666
```

```
Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 5
BD211490
LOCUS       BD211490                 771 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Enzymatically-active recombinant human beta-tryptase and method of
            making same.
ACCESSION   BD211490
VERSION     BD211490.1  GI:33021260
KEYWORDS    JP 2002515254-A/4.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 771)
  AUTHORS   Maffitt,M.A., Niles,A.L. and Frendscho,M.H.
  TITLE     Enzymatically-active recombinant human beta-tryptase and method of
            making same
  JOURNAL   Patent: JP 2002515254-A 4 28-MAY-2002;
            PROMEGA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002515254-A/4
            PD   28-MAY-2002
            PF   29-OCT-1998 JP 2000549745
            PR   15-MAY-1998 US    09/079970
            PI   MARK A MAFFITT,ANDREW L NILES,MARY HAAK FRENDSCHO PC
            C12N15/09,C07K16/40,C12N1/19,C12N9/64,C12P21/08,C12Q1/37// PC
            (C12N9/64,C12R1:645),C12N15/00
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Enzymatically-active recombinant human beta-tryptase and  CC
                    method of making
            CC   same
            FH   Key             Location/Qualifiers
            FT   CDS             7. .753
            FT   misc_signal     7. .18.
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.:              6.53e-122       Length:         771
Score:                  1387.00         Matches:        248
Percent Similarity:     99.60%          Conservative:   0
Best Local Similarity:  99.60%          Mismatches:     1
Query Match:            99.57%          Indels:         0
DB:                     6               Gaps:           0

US-09-598-982-21 (1-249) x BD211490 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 6
AX347857
LOCUS       AX347857                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 24 from Patent WO0198470.
ACCESSION   AX347857
VERSION     AX347857.1  GI:18495664
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 24 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7. .756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22422.1"
                     /db_xref="GI:18495665"
                     /db_xref="REMTREMBL:CAD22422"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN
```

```
Alignment Scores:
Pred. No.:                1.26e-121        Length:          771
Score:                    1384.00          Matches:         247
Percent Similarity:       99.60%           Conservative:    1
Best Local Similarity:    99.20%           Mismatches:      1
Query Match:              99.35%           Indels:          0
DB:                       6                Gaps:            0

US-09-598-982-21 (1-249) x AX347857 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCCGGCGGA 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCTCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 7
AX347859
LOCUS       AX347859                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 26 from Patent WO0198470.
ACCESSION   AX347859
VERSION     AX347859.1  GI:18495666
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 26 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22423.1"
                     /db_xref="GI:18495667"
                     /db_xref="REMTREMBL:CAD22423"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              1.26e-121       Length:         771
Score:                  1384.00         Matches:        247
Percent Similarity:     99.60%          Conservative:   1
Best Local Similarity:  99.20%          Mismatches:     1
Query Match:            99.35%          Indels:         0
DB:                     6               Gaps:           0

US-09-598-982-21 (1-249) x AX347859 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCCGGCGGA 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCACTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 8
AX347838
LOCUS       AX347838                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 5 from Patent WO0198470.
ACCESSION   AX347838
VERSION     AX347838.1  GI:18495645
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 5 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22417.1"
                     /db_xref="GI:18495646"
                     /db_xref="REMTREMBL:CAD22417"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              1.94e-121       Length:         771
Score:                  1382.00         Matches:        247
Percent Similarity:     99.20%          Conservative:   0
Best Local Similarity:  99.20%          Mismatches:     2
Query Match:            99.21%          Indels:         0
DB:                     6               Gaps:           0

US-09-598-982-21 (1-249) x AX347838 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246
```

```
Qy          81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy         101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
               |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db         307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy         121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy         141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy         161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy         181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy         201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy         221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy         241 IleHisHisTyrValProLysLysPro 249
               |||||||||||||||||||||||||||
Db         727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 9
AX347855
LOCUS       AX347855                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 22 from Patent WO0198470.
ACCESSION   AX347855
VERSION     AX347855.1  GI:18495662
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 22 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22421.1"
                     /db_xref="GI:18495663"
                     /db_xref="REMTREMBL:CAD22421"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAAIALLELE
                     EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
```

```
Pred. No.:              3.73e-121       Length:          771
Score:                  1379.00         Matches:         247
Percent Similarity:     99.20%          Conservative:    0
Best Local Similarity:  99.20%          Mismatches:      2
Query Match:            98.99%          Indels:          0
DB:                     6               Gaps:            0

US-09-598-982-21 (1-249) x AX347855 (1-771)

Qy           1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy          21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy          41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
               |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy          61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy          81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
               ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db         247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTGCTGGAG 306

Qy         101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy         121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy         141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy         161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy         181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy         201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy         221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy         241 IleHisHisTyrValProLysLysPro 249
               |||||||||||||||||||||||||||
Db         727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 10
AX347873
LOCUS       AX347873                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 40 from Patent WO0198470.
ACCESSION   AX347873
VERSION     AX347873.1  GI:18495680
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
```

```
 AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
 TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
           methods of making same
 JOURNAL   Patent: WO 0198470-A 40 27-DEC-2001;
           PROMEGA CORPORATION (US)
FEATURES            Location/Qualifiers
     source         1. .771
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"
     CDS            <1. .756
                    /note="unnamed protein product"
                    /codon_start=1
                    /protein_id="CAD22430.1"
                    /db_xref="GI:18495681"
                    /db_xref="REMTREMBL:CAD22430"
                    /translation="GPLEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQW
                    VLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLE
                    LEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIME
                    NHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVV
                    SWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              3.73e-121      Length:        771
Score:                  1379.00        Matches:       246
Percent Similarity:     99.20%         Conservative:  1
Best Local Similarity:  98.80%         Mismatches:    2
Query Match:            98.99%         Indels:        0
DB:                     6              Gaps:          0

US-09-598-982-21 (1-249) x AX347873 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCCGGCGGA 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCTCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 11
AX347875
LOCUS       AX347875                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 42 from Patent WO0198470.
ACCESSION   AX347875
VERSION     AX347875.1  GI:18495682
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 42 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7. .756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22431.1"
                     /db_xref="GI:18495683"
                     /db_xref="REMTREMBL:CAD22431"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              3.73e-121      Length:        771
Score:                  1379.00        Matches:       246
Percent Similarity:     99.20%         Conservative:  1
Best Local Similarity:  98.80%         Mismatches:    2
Query Match:            98.99%         Indels:        0
DB:                     6              Gaps:          0

US-09-598-982-21 (1-249) x AX347875 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246
```

```
Qy          81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy         101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
               |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db         307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy         121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy         141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy         161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy         181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
               |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db         547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCCGGCGGA 606

Qy         201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         607 CCACTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy         221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy         241 IleHisHisTyrValProLysLysPro 249
               |||||||||||||||||||||||||||
Db         727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 12
AX347861
LOCUS       AX347861                 735 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 28 from Patent WO0198470.
ACCESSION   AX347861
VERSION     AX347861.1  GI:18495668
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 28 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .735
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <1. .>735
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22424.1"
                     /db_xref="GI:18495669"
                     /db_xref="REMTREMBL:CAD22424"
                     /translation="IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAA
                     CVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVK
                     VSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDA
                     KYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGC
                     AQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:              1.05e-120      Length:        735
```

```
Score:                      1374.00       Matches:        245
Percent Similarity:         100.00%       Conservative:   0
Best Local Similarity:      100.00%       Mismatches:     0
Query Match:                98.64%        Indels:         0
DB:                         6             Gaps:           0

US-09-598-982-21 (1-249) x AX347861 (1-735)

Qy           5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy          25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy          45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy          65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy          85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy         105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy         125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy         145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy         165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy         185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy         205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy         225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy         245 ValProLysLysPro 249
               |||||||||||||||
Db         721 GTCCCCAAAAAGCCG 735

RESULT 13
AX347871
LOCUS       AX347871                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 38 from Patent WO0198470.
ACCESSION   AX347871
VERSION     AX347871.1  GI:18495678
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
```

```
TITLE       Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
JOURNAL     Patent: WO 0198470-A 38 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7. .756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22429.1"
                     /db_xref="GI:18495679"
                     /db_xref="REMTREMBL:CAD22429"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAAIALLELE
                     EPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:                1.11e-120     Length:          771
Score:                    1374.00       Matches:         246
Percent Similarity:       98.80%        Conservative:    0
Best Local Similarity:    98.80%        Mismatches:      3
Query Match:              98.64%        Indels:          0
DB:                       6             Gaps:            0

US-09-598-982-21 (1-249) x AX347871 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 14
AX347877
LOCUS       AX347877                  735 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 44 from Patent WO0198470.
ACCESSION   AX347877
VERSION     AX347877.1  GI:18495684
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 44 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .735
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <1. .>735
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22432.1"
                     /db_xref="GI:18495685"
                     /db_xref="REMTREMBL:CAD22432"
                     /translation="IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAA
                     CVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVN
                     VSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDA
                     KYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGC
                     AQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

Alignment Scores:
Pred. No.:                3.12e-120     Length:          735
Score:                    1369.00       Matches:         244
Percent Similarity:       99.59%        Conservative:    0
Best Local Similarity:    99.59%        Mismatches:      1
Query Match:              98.28%        Indels:          0
DB:                       6             Gaps:            0

US-09-598-982-21 (1-249) x AX347877 (1-735)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy         105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
                 |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy         125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy         145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy         165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy         185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy         205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy         225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy         245 ValProLysLysPro 249
                 |||||||||||||||
Db         721 GTCCCCAAAAAGCCG 735

RESULT 15
AR165109
LOCUS       AR165109                 735 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 1 from patent US 6274366.
ACCESSION   AR165109
VERSION     AR165109.1  GI:16238522
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 735)
  AUTHORS   Maffitt,M.A., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Enzymatically-active recombinant human .beta.-tryptase and method
            of making same
  JOURNAL   Patent: US 6274366-A 1 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1..735
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

Alignment Scores:
Pred. No.:              3.88e-120       Length:         735
Score:                  1368.00         Matches:        244
Percent Similarity:     99.59%          Conservative:   0
Best Local Similarity:  99.59%          Mismatches:     1
Query Match:            98.21%          Indels:         0
DB:                     6               Gaps:           0

US-09-598-982-21 (1-249) x AR165109 (1-735)

Qy           5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy          25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120
```

```
Qy          45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
                 |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy          65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy          85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy         105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy         125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy         145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy         165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy         185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy         205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy         225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy         245 ValProLysLysPro 249
                 |||||||||||||||
Db         721 GTCCCCAAAAAGCCG 735

Search completed: July 22, 2004, 17:22:57
Job time : 4805 secs
```

```
                 GenCore version 5.1.6
         Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:          July 22, 2004, 15:38:42 ; Search time 86 Seconds
                                             (without alignments)
                                             1606.776 Million cell updates/sec

Title:           US-09-598-982-21
Perfect score:   1393
Sequence:        1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:      1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_20072004_064410_6521/app_query.fasta_1.391
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982_@CGN_1_1_128_@runat_20072004_064410_6521 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       Issued_Patents_NA:*
                 1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                 2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                 3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                 4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                 5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                 6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                  %
Result          Query
  No.   Score   Match Length DB  ID                      Description
------------------------------------------------------------------------------
    1    1387    99.6    771  3  US-09-079-970A-4        Sequence 4, Appli
    2    1368    98.2    735  3  US-09-079-970A-1        Sequence 1, Appli
    3    1368    98.2   1128  2  US-09-016-366A-20       Sequence 20, Appl
    4    1368    98.2   1128  2  US-08-978-404B-15       Sequence 15, Appl
    5    1363    97.8   1137  2  US-09-016-366A-18       Sequence 18, Appl
    6    1363    97.8   1137  2  US-08-978-404B-13       Sequence 13, Appl
    7    1344    96.5   1081  2  US-09-016-366A-22       Sequence 22, Appl
    8    1344    96.5   1081  2  US-08-978-404B-17       Sequence 17, Appl
    9    1274    91.5   1154  2  US-09-016-366A-16       Sequence 16, Appl
   10    1274    91.5   1154  2  US-08-978-404B-11       Sequence 11, Appl
   11    1091    78.3   1108  2  US-09-016-366A-14       Sequence 14, Appl
   12    1091    78.3   1108  2  US-08-978-404B-20       Sequence 20, Appl
```

```
   13    1075    77.2   1219  2  US-08-978-404B-7        Sequence 7, Appli
   14    1054    75.7   1031  2  US-08-978-404B-1        Sequence 1, Appli
   15    1051    75.4   1097  2  US-08-978-404B-4        Sequence 4, Appli
   16    1051    75.4   1103  2  US-09-016-366A-24       Sequence 24, Appl
   17  1000.5    71.8   2259  2  US-08-845-998-3         Sequence 3, Appli
   18  1000.5    71.8   2259  3  US-09-206-537-3         Sequence 3, Appli
   19  1000.5    71.8   2259  3  US-09-430-854-3         Sequence 3, Appli
   20     979    70.3   3757  2  US-09-016-366A-13       Sequence 13, Appl
   21     979    70.3   3757  2  US-08-978-404B-19       Sequence 19, Appl
   22   961.5    69.0   2218  2  US-08-845-998-5         Sequence 5, Appli
   23   961.5    69.0   2218  3  US-09-206-537-5         Sequence 5, Appli
   24   961.5    69.0   2218  3  US-09-430-854-5         Sequence 5, Appli
   25   836.5    60.1   2397  2  US-08-978-404B-2        Sequence 2, Appli
   26   684.5    49.1   1095  2  US-08-978-404B-9        Sequence 9, Appli
   27   563.5    40.5    980  4  US-09-023-942A-30       Sequence 30, Appl
   28   563.5    40.5   1110  4  US-09-386-653A-1        Sequence 1, Appli
   29   561.5    40.3   1212  4  US-09-620-312D-431      Sequence 431, App
   30   554.5    39.8   1130  4  US-09-386-653A-8        Sequence 8, Appli
   31     552    39.6    933  4  US-09-023-942A-29       Sequence 29, Appl
   32   536.5    38.5   1613  4  US-09-387-375-1         Sequence 1, Appli
   33   527.5    37.9   1130  4  US-09-387-375-8         Sequence 8, Appli
   34     507    36.4   1142  4  US-09-386-642-8         Sequence 8, Appli
   35     507    36.4   1169  4  US-09-386-642-7         Sequence 7, Appli
   36   501.5    36.0   1165  4  US-09-023-942A-28       Sequence 28, Appl
   37   501.5    36.0   1378  4  US-09-907-794A-262      Sequence 262, App
   38   501.5    36.0   1378  4  US-09-905-125A-262      Sequence 262, App
   39   501.5    36.0   1378  4  US-09-902-775A-262      Sequence 262, App
   40   501.5    36.0   1430  4  US-09-386-629-1         Sequence 1, Appli
   41   499.5    35.9   1094  4  US-09-023-942A-3        Sequence 3, Appli
   42   493.5    35.4   1081  3  US-09-008-271A-15       Sequence 15, Appl
   43   493.5    35.4   1100  4  US-09-907-794A-256      Sequence 256, App
   44   493.5    35.4   1100  4  US-09-905-125A-256      Sequence 256, App
   45   493.5    35.4   1100  4  US-09-902-775A-256      Sequence 256, App

                                  ALIGNMENTS

RESULT 1
US-09-079-970A-4
; Sequence 4, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  771 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  7..753
;    FEATURE:
;      NAME/KEY:  misc_signal
;      LOCATION:  7..18
US-09-079-970A-4

Alignment Scores:
Pred. No.:                1.71e-144     Length:         771
Score:                    1387.00       Matches:        248
Percent Similarity:       99.60%        Conservative:   0
Best Local Similarity:    99.60%        Mismatches:     1
Query Match:              99.57%        Indels:         0
DB:                       3             Gaps:           0

US-09-598-982-21 (1-249) x US-09-079-970A-4 (1-771)

Qy            1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy           21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy           41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
                |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db          127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy           61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy           81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy          101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy          121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy          141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy          161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy          181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy          201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy          221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy          241 IleHisHisTyrValProLysLysPro 249
                |||||||||||||||||||||||||||
Db          727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 2
US-09-079-970A-1
; Sequence 1, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  735 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..735
US-09-079-970A-1

Alignment Scores:
Pred. No.:                2.04e-142     Length:         735
Score:                    1368.00       Matches:        244
Percent Similarity:       99.59%        Conservative:   0
Best Local Similarity:    99.59%        Mismatches:     1
Query Match:              98.21%        Indels:         0
DB:                       3             Gaps:           0

US-09-598-982-21 (1-249) x US-09-079-970A-1 (1-735)

Qy            5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy           25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120
```

```
Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        721 GTCCCCAAAAAGCCG 735

RESULT 3
US-09-016-366A-20
; Sequence 20, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-20

Alignment Scores:
Pred. No.:              3.84e-142     Length:         1128
Score:                  1368.00       Matches:        244
Percent Similarity:     99.59%        Conservative:   0
Best Local Similarity:  99.59%        Mismatches:     1
Query Match:            98.21%        Indels:         0
DB:                     2             Gaps:           0

US-09-598-982-21 (1-249) x US-09-016-366A-20 (1-1128)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         89 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 148

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 208

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        209 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 268

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 328

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 388

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 448

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 508

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 568

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 628

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 688

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 748

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          749 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 808

Qy          245 ValProLysLysPro 249
                |||||||||||||||
Db          809 GTCCCCAAAAAGCCG 823

RESULT 4
US-08-978-404B-15
; Sequence 15, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-15

 Alignment Scores:
 Pred. No.:              3.84e-142      Length:          1128
 Score:                  1368.00        Matches:         244
 Percent Similarity:     99.59%         Conservative:    0
 Best Local Similarity:  99.59%         Mismatches:      1
 Query Match:            98.21%         Indels:          0
 DB:                     2              Gaps:            0

US-09-598-982-21 (1-249) x US-08-978-404B-15 (1-1128)

Qy            5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           89 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 148

Qy           25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          149 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 208

Qy           45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
                |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db          209 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 268
```

```
Qy           65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          269 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 328

Qy           85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          329 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 388

Qy          105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          389 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 448

Qy          125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          449 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 508

Qy          145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          509 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 568

Qy          165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          569 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 628

Qy          185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          629 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 688

Qy          205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          689 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 748

Qy          225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          749 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 808

Qy          245 ValProLysLysPro 249
                |||||||||||||||
Db          809 GTCCCCAAAAAGCCG 823

RESULT 5
US-09-016-366A-18
; Sequence 18, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-18

Alignment Scores:
Pred. No.:             1.39e-141    Length:        1137
Score:                 1363.00      Matches:       243
Percent Similarity:    99.18%       Conservative:  0
Best Local Similarity: 99.18%       Mismatches:    2
Query Match:           97.85%       Indels:        0
DB:                    2            Gaps:          0

US-09-598-982-21 (1-249) x US-09-016-366A-18 (1-1137)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 146

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 206

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 266

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 326

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 386

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 446

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 506

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 566

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 626

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 686

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 746

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 806

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        807 GTCCCCAAAAAGCCG 821

RESULT 6
US-08-978-404B-13
; Sequence 13, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-13

Alignment Scores:
Pred. No.:             1.39e-141    Length:        1137
Score:                 1363.00      Matches:       243
Percent Similarity:    99.18%       Conservative:  0
Best Local Similarity: 99.18%       Mismatches:    2
Query Match:           97.85%       Indels:        0
DB:                    2            Gaps:          0

US-09-598-982-21 (1-249) x US-08-978-404B-13 (1-1137)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 146

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 206

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        207 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 266

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 326

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 386

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 446

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 506

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 566

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 626

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 686

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 746

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 806

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        807 GTCCCCAAAAAGCCG 821

RESULT 7
US-09-016-366A-22
; Sequence 22, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-22

Alignment Scores:
Pred. No.:                  1.65e-139       Length:          1081
Score:                      1344.00         Matches:         241
Percent Similarity:         98.37%          Conservative:    0
Best Local Similarity:      98.37%          Mismatches:      4
Query Match:                96.48%          Indels:          0
DB:                         2               Gaps:            0

US-09-598-982-21 (1-249) x US-09-016-366A-22 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 8
US-08-978-404B-17
; Sequence 17, Application US/08978404B
```

```
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-17

Alignment Scores:
Pred. No.:              1.65e-139     Length:          1081
Score:                  1344.00       Matches:         241
Percent Similarity:     98.37%        Conservative:    0
Best Local Similarity:  98.37%        Mismatches:      4
Query Match:            96.48%        Indels:          0
DB:                     2             Gaps:            0

US-09-598-982-21 (1-249) x US-08-978-404B-17 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 9
US-09-016-366A-16
; Sequence 16, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1154 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
```

```
;    MOLECULE TYPE:  cDNA
US-09-016-366A-16

Alignment Scores:
Pred. No.:                 1.05e-131     Length:         1154
Score:                     1274.00       Matches:        226
Percent Similarity:        94.69%        Conservative:   6
Best Local Similarity:     92.24%        Mismatches:     13
Query Match:               91.46%        Indels:         0
DB:                        2             Gaps:           0

US-09-598-982-21 (1-249) x US-09-016-366A-16 (1-1154)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 167

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 227

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   |||:::||||||||||||||||||||||||   |||||||||||||||
Db        228 ACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTG 287

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGCAC 347

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db        348 CCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCC 407

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::||||||   ||||||||||||   |||||||||||||||||||||||||||
Db        408 GTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCC 467

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        468 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCA 527

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 CCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 587

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        588 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTG 647

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||::::::||||||||||||:::|||||||||||||||||||||||||||
Db        648 TGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGCCCCCTGGTGTGC 707

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        708 AAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCCAG 767

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        768 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 827

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        828 GTCCCCAAAAAGCCG 842

RESULT 10
US-08-978-404B-11
; Sequence 11, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1154 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-11

Alignment Scores:
Pred. No.:                 1.05e-131     Length:         1154
Score:                     1274.00       Matches:        226
Percent Similarity:        94.69%        Conservative:   6
Best Local Similarity:     92.24%        Mismatches:     13
Query Match:               91.46%        Indels:         0
DB:                        2             Gaps:           0

US-09-598-982-21 (1-249) x US-08-978-404B-11 (1-1154)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 167

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 227

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   |||:::||||||||||||||||||||||||   |||||||||||||||
Db        228 ACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTG 287

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGCAC 347

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db        348 CCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCC 407

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::||||||   ||||||||||||   |||||||||||||||||||||||||||
Db        408 GTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCC 467

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        468 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCA 527
```

```
Qy           145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           528 CCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 587

Qy           165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
                 ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db           588 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTG 647

Qy           185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
                 ||||||||||||::::::||||||||||||:::|||||||||||||||||||||||||||
Db           648 TGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGCCCCTGGTGTGC 707

Qy           205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
                 ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db           708 AAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCCAG 767

Qy           225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           768 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 827

Qy           245 ValProLysLysPro 249
                 |||||||||||||||
Db           828 GTCCCCAAAAAGCCG 842

RESULT 11
US-09-016-366A-14
; Sequence 14, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-14

Alignment Scores:
Pred. No.:              1.91e-111      Length:         1108
```

```
Score:                     1091.00         Matches:         190
Percent Similarity:        86.01%          Conservative:    19
Best Local Similarity:     78.19%          Mismatches:      34
Query Match:               78.32%          Indels:          0
DB:                        2               Gaps:            0

US-09-598-982-21 (1-249) x US-09-016-366A-14 (1-1108)

Qy             5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
                 ||||||||||||   ||||||      ||||||||||||||||||||||||||||||
Db           127 ATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGATTT 186

Qy            25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                          ||||||:::||||||||||||||||||||||||||||||||||||||||||
Db           187 AAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGCTC 246

Qy            45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
                 |||||||||   ||||||||||||   :::|||               ||||||||||||
Db           247 ACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGCTT 306

Qy            65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
                 |||||||||:::|||||||||   ||||||||||||   ::::::||||||:::||||||
Db           307 CGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGCAC 366

Qy            85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
                 |||   :::|||||||||:::   |||||||||:::||||||||||||||||||   |||
Db           367 CCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCCCT 426

Qy           105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
                 |||   ||||||:::|||:::|||   ::::::|||||||||||||||||||||||||||
Db           427 GTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCCCC 486

Qy           125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                 ||||||      ||||||||||||||||||||||||:::||||||||||||   ||||||
Db           487 CCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCCCA 546

Qy           145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
                 ||||||:::|||||||||||||||||||||||||||:::||||||   :::||||||
Db           547 CCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACCGG 606

Qy           165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
                 |||||||||   |||   |||||||||||||||      ||||||   |||   ||||||
Db           607 AAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGCTG 666

Qy           185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           667 TGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGCCACTGGTCTGC 726

Qy           205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
                 ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db           727 AAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGGTGAGGGCTGCGCACAG 786

Qy           225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
                 ||||||:::|||||||||||||||||||||||||||||||||||||||||||||   |||
Db           787 CCCAACAAGCCTGGCATCTACACCCGGGTGACATACTACTTAGACTGGATCCACCGCTAT 846

Qy           245 ValProLys 247
                 ||||||:::
Db           847 GTCCCTGAG 855

RESULT 12
US-08-978-404B-20
; Sequence 20, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
```

```
;       STATE:  MA
;       COUNTRY:  U.S.A.
;       ZIP:  02210-2211
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/978,404B
;       FILING DATE:  25-NOV-97
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/032,354
;       FILING DATE:  04-DEC-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Plumer, Elizabeth R.
;       REGISTRATION NUMBER:  36,637
;       REFERENCE/DOCKET NUMBER:  B0801/7090
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-720-3500
;       TELEFAX:  617-720-2441
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  20:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1108 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-978-404B-20

Alignment Scores:
Pred. No.:               1.91e-111     Length:          1108
Score:                   1091.00       Matches:         190
Percent Similarity:      86.01%        Conservative:    19
Best Local Similarity:   78.19%        Mismatches:      34
Query Match:             78.32%        Indels:          0
DB:                      2             Gaps:            0

US-09-598-982-21 (1-249) x US-08-978-404B-20 (1-1108)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||   ||||||      ||||||||||||||||||||||||||||||
Db        127 ATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGATTT 186

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       ||||||:::|||||||||||||||||||||||||||||||||||||||
Db        187 AAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGCTC 246

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||   :::|||               ||||||||||||
Db        247 ACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGCTT 306

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              |||||||||:::|||||||||   ||||||||||||   ::::::||||||:::||||||
Db        307 CGTGAGCAGTATCTATACTATGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGCAC 366

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              |||   :::|||||||||:::   |||||||||:::||||||||||||||||||   |||
Db        367 CCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCCCT 426

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||:::|||:::|||   ::::::|||||||||||||||||||||||||||
Db        427 GTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCCCC 486

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||      ||||||||||||||||||||||||:::||||||||||||   ||||||
Db        487 CCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCCCA 546

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||:::|||||||||||||||||||||||||||:::||||||   :::||||||
Db        547 CCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACCGG 606
```

```
Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              |||||||||   |||   |||||||||||||||      ||||||   |||   ||||||
Db        607 AAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGCTG 666

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 TGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGCCACTGGTCTGC 726

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        727 AAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGGTGAGGGCTGCGCACAG 786

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||:::|||||||||||||||||||||||||||||||||||||||||||   |||
Db        787 CCCAACAAGCCTGGCATCTACACCCGGGTGACATACTACTTAGACTGGATCCACCGCTAT 846

Qy        245 ValProLys 247
              ||||||:::
Db        847 GTCCCTGAG 855

RESULT 13
US-08-978-404B-7
; Sequence 7, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1219 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-7

Alignment Scores:
Pred. No.:               1.31e-109     Length:          1219
Score:                   1075.00       Matches:         188
Percent Similarity:      84.77%        Conservative:    18
Best Local Similarity:   77.37%        Mismatches:      37
Query Match:             77.17%        Indels:          0
DB:                      2             Gaps:            0

US-09-598-982-21 (1-249) x US-08-978-404B-7 (1-1219)
```

```
Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||   :::|||||||||||||||||||||||||||
Db        273 ATTGTTGGGGACAGGAGGCACCTGGGAACAAGTGGCCCTGGCAGGTGAGCCTTCGTGCC 332

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              :::      ||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        333 AATGAAACCTACTGGAGGCATTTCTGCGGCGGCTCCCTCATCCACCCACAGTGGGTGCTC 392

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||   :::   |||         :::||||||||||||
Db        393 ACCGCGGCACACTGTGTGGGACCGACTATTGCTGATCCCAACAAGGTCAGAGTACAGCTT 452

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              |||:::|||:::|||||||||   |||   ||||||   |||||||||||||||   |||
Db        453 CGAAAGCAGTACCTCTATTACCACGACCACCTGCTGGCTGTGAGCCGGATCATCACACAC 512

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              |||   ||||||      |||   |||||||||||||||||||||||||||:::   |||
Db        513 CCGACATTCTATGCCACCCAGAATGGGGCGGACATCGCCCTACTTGAGCTCAAGAACCCT 572

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::|||||||||||||||   |||:::|||||||||||||||||||||||||||
Db        573 GTAAACATTTCCAGCCATGTCCACCCCGTCTCCCTGCCTCCTGCCTCAGAGACCTTCCCC 632

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                 |||      |||||||||||||||||||||::::::|||||||||      ||||||
Db        633 TCAGGAACATTGTGCTGGGTGACAGGCTGGGGAAACATCGACAATGATGTGAGCCTGCCA 692

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||:::|||:::||||||::::::||||||   :::||||||
Db        693 CCGCCATTTCCCTTGAAGGAGGTGCAAGTTCCCGTCGTGGAAAACCAGCTTTGTGACCTG 752

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              |||||||||   |||   ||||||||||||::::::   |||||||||||||||||||||
Db        753 AAGTATCACAAAGGTGTCTACACAGGGGACAACATCCACATTGTCCGAGACGACATGCTG 812

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||         |||||||||||||||||||||||||||||||||||||||
Db        813 TGTGCTGGGAACGAAGGACACGACTCCTGCCAGGGTGACTCCGGAGGACCTCTGGTCTGC 872

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        873 AAGGTAAACGGTACCTGGCTGCAGGCAGGTGTGGTCAGCTGGGGTGAGGGCTGTGCTCTG 932

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        933 CCCAACAGGCCTGGCATCTACACTCGGGTCACCTATTACTTGGACTGGATCCACCGCTAT 992

Qy        245 ValProLys 247
              |||||||||
Db        993 GTCCCCAAG 1001

RESULT 14
US-08-978-404B-1
; Sequence 1, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
```

```
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1031 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-1

Alignment Scores:
Pred. No.:              2.17e-107        Length:         1031
Score:                  1054.00          Matches:        186
Percent Similarity:     83.54%           Conservative:   17
Best Local Similarity:  76.54%           Mismatches:     40
Query Match:            75.66%           Indels:         0
DB:                     2                Gaps:           0

US-09-598-982-21 (1-249) x US-08-978-404B-1 (1-1031)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              |||||||||||||||||||||      :::|||||||||||||||||||||||||||
Db        111 ATTGTGGGGGGACAGGAGGCACATGGGAACAAGTGGCCCTGGCAGGTGAGCCTGCGTGCC 170

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              :::      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        171 AATGACACCTACTGGATGCATTTCTGCGGTGGCTCCCTCATCCACCCACAGTGGGTGCTC 230

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||   |||         :::||||||||||||
Db        231 ACTGCGGCACACTGTGTGGGACCGGATGTTGCTGACCCCAACAAGGTCAGAGTACAGCTC 290

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              |||:::|||:::|||||||||   |||   |||:::   ||||||:::||||||   |||
Db        291 CGTAAGCAGTACCTCTATTACCATGACCACCTGATGACTGTGAGCCAGATCATCACACAC 350

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              |||   ||||||      |||   |||||||||||||||||||||:::|||      |||
Db        351 CCCGACTTCTACATCGTCCAGGATGGGGCAGACATTGCCCTGCTGAAACTCACAAACCCT 410

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::|||   :::||||||   |||   |||||||||||||||||||||||||||
Db        411 GTGAACATTTCTGACTATGTCCACCCTGTCCCCCTACCTCCTGCCTCAGAGACCTTCCCC 470

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                 |||      |||||||||||||||||||||::::::||||||         ||||||
Db        471 TCAGGAACGTTGTGCTGGGTGACAGGCTGGGGTAACATCGACAATGGTGTAAACCTGCCG 530

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||:::|||:::|||||||||:::|||||||||:::||||||
Db        531 CCACCATTTCCTTTGAAGGAGGTGCAAGTTCCCATTATAGAAAACCACCTTTGTGACTTG 590

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              |||||||||   |||      |||||||||:::|||   |||||||||||||||||||||
Db        591 AAGTATCACAAAGGTCTCATCACAGGTGACAATGTCCACATTGTCCGAGATGACATGCTG 650

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||         |||||||||||||||||||||||||||||||||||||||
Db        651 TGTGCTGGGAATGAAGGACATGACTCCTGCCAGGGCGACTCCGGAGGACCTCTGGTCTGC 710
```

```
Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db        711 AAGGTAGAAGACACCTGGCTGCAGGCAGGCGTGGTCAGCTGGGGTGAGGGCTGTGCACAG 770

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        771 CCCAACAGGCCTGGCATCTACACCCGGGTCACCTATTACTTGGACTGGATCCACCACTAT 830

Qy        245 ValProLys 247
              |||||||||
Db        831 GTCCCCAAG 839

RESULT 15
US-08-978-404B-4
; Sequence 4, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1097 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-4

Alignment Scores:
Pred. No.:              5.12e-107      Length:         1097
Score:                  1051.00        Matches:        185
Percent Similarity:     84.43%         Conservative:   21
Best Local Similarity:  75.82%         Mismatches:     38
Query Match:            75.45%         Indels:         0
DB:                     2              Gaps:           0

US-09-598-982-21 (1-249) x US-08-978-404B-4 (1-1097)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||:::||||||      ||||||||||||||||||||||||||||||
Db        113 ATTGTGGGAGGACGAGAGGCTTCTGAAAGTAAGTGGCCCTGGCAGGTGAGCCTGAGATTT 172

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       :::||||||||||||||||||||||||||||||||||||||||||||||||
Db        173 AAATTCAGCTTCTGGATGCATTTCTGTGGCGGCTCCCTCATTCACCCACAGTGGGTGCTC 232

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   |||||||||      :::|||               ||||||||||||
Db        233 ACTGCGGCACACTGTGTGGGACTGCACATCAAAAGCCCAGAGCTCTTCCGTGTACAGCTT 292

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              |||||||||:::|||||||||   ||||||||||||   |||:::|||   :::||||||
Db        293 CGTGAGCAGTATCTATACTATGCGGACCAGCTACTGACTGTGAACCGGACCGTTGTGCAC 352

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              |||   :::||||||   :::   ||||||||||||||||||||||||||||||   |||
Db        353 CCCCACTACTACACAGTCGAGGATGGGCAGACATTGCCCTGCTGGAGCTTGAGATCCCT 412

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||:::|||:::|||   ::::::|||||||||||||||||||||||||||
Db        413 GTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCCCC 472

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                 |||      ||||||||||||||||||||||||:::|||:::||||||   |||
Db        473 TCGGGGACTTCTTGCTGGGTAACAGGCTGGGGCGACATTGATAGTGACGAGCCTCTCCTG 532

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||:::|||||||||||||||||||||||||||:::||||||   :::||||||
Db        533 CCACCTTATCCTCTGAAGCAAGTGAAGGTCCCCATTGTGGAAAACAGCCTGTGTGATCGG 592

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              |||||||||   |||   ||||||||||||||||||   ||||||:::|||   ||||||
Db        593 AAGTACCACACTGGCCTCTACACAGGAGATGATGTTCCCATTGTCCAGGATGGCATGCTG 652

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db        653 TGTGCTGGAAATACCAGGAGCGACTCCTGCCAGGGAGACTCAGGGGGCCCACTGGTCTGC 712

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||:::
Db        713 AAAGTGAAGGGTACCTGGCTGCAAGCAGGAGTGGTCAGCTGGGGTGAGGGCTGCGCAGAG 772

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
                 |||||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db        773 GCCAATCGTCCTGGCATTTACACCCGGGTGACGTACTACCTGGACTGGATTCACCGCTAT 832

Qy        245 ValProLysLys 248
              ||||||::::::
Db        833 GTCCCTCAGCGT 844


Search completed: July 22, 2004, 18:16:00
Job time : 92 secs
```

```
                          GenCore version 5.1.6
                 Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:          July 22, 2004, 15:26:22 ; Search time 474 Seconds
                                           (without alignments)
                                           2231.648 Million cell updates/sec

Title:           US-09-598-982-21
Perfect score:   1393
Sequence:        1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:       6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_20072004_064409_6476/app_query.fasta_1.391
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982_@CGN_1_1_885_@runat_20072004_064409_6476 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       N_Geneseq_29Jan04:*
                 1:  geneseqn1980s:*
                 2:  geneseqn1990s:*
                 3:  geneseqn2000s:*
                 4:  geneseqn2001as:*
                 5:  geneseqn2001bs:*
                 6:  geneseqn2002s:*
                 7:  geneseqn2003as:*
                 8:  geneseqn2003bs:*
                 9:  geneseqn2003cs:*
                 10:  geneseqn2004s:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                  SUMMARIES

                      %
Result               Query
  No.    Score       Match  Length DB  ID                   Description
-----------------------------------------------------------------------------
    1     1393       100.0     771  6  AAS20775             Aas20775 DNA encod
    2     1388        99.6     771  6  AAS20783             Aas20783 DNA encod
    3     1387        99.6     771  3  AAZ40175             Aaz40175 Human bet
    4     1387        99.6     771  6  AAS20765             Aas20765 DNA encod
    5     1384        99.4     771  6  AAS20778             Aas20778 DNA encod
    6     1384        99.4     771  6  AAS20777             Aas20777 DNA encod
    7     1382        99.2     771  6  AAS20763             Aas20763 DNA encod
    8     1379        99.0     771  6  AAS20786             Aas20786 DNA encod
```

```
    9     1379        99.0     771  6  AAS20785             Aas20785 DNA encod
   10     1379        99.0     771  6  AAS20776             Aas20776 DNA encod
   11     1374        98.6     735  6  AAS20779             Aas20779 DNA encod
   12     1374        98.6     771  6  AAS20784             Aas20784 DNA encod
   13     1369        98.3     735  6  AAS20787             Aas20787 DNA encod
   14     1368        98.2     735  3  AAZ40172             Aaz40172 Human bet
   15     1368        98.2     735  6  AAS20766             Aas20766 DNA encod
   16     1368        98.2    1128  2  AAV44330             Aav44330 Human mas
   17     1368        98.2    1128  2  AAV42712             Aav42712 Human mas
   18     1368        98.2    1143  3  AAA34955             Aaa34955 Human ade
   19     1368        98.2    1143  3  AAF21077             Aaf21077 Human low
   20     1368        98.2    1143  7  ABZ96771             Abz96771 Human nuc
   21     1368        98.2    1145  3  AAA34956             Aaa34956 Human ade
   22     1368        98.2    1145  3  AAF21078             Aaf21078 Human low
   23     1368        98.2    1145  7  ABZ96772             Abz96772 Human nuc
   24     1368        98.2   17133  3  AAA34962             Aaa34962 Human ade
   25     1368        98.2   17133  3  AAF21084             Aaf21084 Human low
   26     1368        98.2   17133  7  ABZ96778             Abz96778 Human nuc
   27     1365        98.0     735  6  AAS20782             Aas20782 DNA encod
   28     1365        98.0     735  6  AAS20781             Aas20781 DNA encod
   29     1363        97.8     735  6  AAS20760             Aas20760 DNA encod
   30     1363        97.8    1137  2  AAV44329             Aav44329 Human mas
   31     1363        97.8    1137  2  AAV42711             Aav42711 Human mas
   32     1363        97.8    1137  3  AAA34957             Aaa34957 Human ade
   33     1363        97.8    1137  3  AAF21079             Aaf21079 Human low
   34     1363        97.8    1137  7  ABZ96773             Abz96773 Human nuc
   35     1360        97.6     735  6  AAS20789             Aas20789 DNA encod
   36     1360        97.6     735  6  AAS20790             Aas20790 DNA encod
   37     1360        97.6     735  6  AAS20780             Aas20780 DNA encod
   38     1355        97.3     735  6  AAS20788             Aas20788 DNA encod
   39     1344        96.5    1081  2  AAV44331             Aav44331 Human mas
   40     1344        96.5    1081  2  AAV42713             Aav42713 Human mas
   41     1344        96.5    1081  3  AAA34960             Aaa34960 Human ade
   42     1344        96.5    1081  3  AAF21082             Aaf21082 Human low
   43     1344        96.5    1081  6  ABL62377             Abl62377 Colon ade
   44     1344        96.5    1081  6  ABL66816             Abl66816 Lung canc
   45     1344        96.5    1081  6  ABL61818             Abl61818 Colon ade

                                  ALIGNMENTS

RESULT 1
AAS20775
ID   AAS20775 standard; DNA; 771 BP.
XX
AC   AAS20775;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant H44A #1.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12009.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
```

```
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 83-84; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 147 A; 256 C; 244 G; 124 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              2.04e-121     Length:          771
Score:                  1393.00       Matches:         249
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            100.00%       Indels:          0
DB:                     6             Gaps:            0

US-09-598-982-21 (1-249) x AAS20775 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 2
AAS20783
ID   AAS20783 standard; DNA; 771 BP.
XX
AC   AAS20783;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant H44A #3.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12017.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 104-105; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 147 A; 257 C; 243 G; 124 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              6.02e-121     Length:          771
```

```
Score:                   1388.00         Matches:       248
Percent Similarity:      99.60%          Conservative:  0
Best Local Similarity:   99.60%          Mismatches:    1
Query Match:             99.64%          Indels:        0
DB:                      6               Gaps:          0

US-09-598-982-21 (1-249) x AAS20783 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 3
AAZ40175
ID   AAZ40175 standard; DNA; 771 BP.
XX
AC   AAZ40175;
XX
DT   18-FEB-2000  (first entry)
XX
DE   Human beta-tryptase coding sequence.
XX
KW   Beta-tryptase; human; DNA expression construct; protein production;
KW   combinatorial library screening; X ray crystallography; antigen;
KW   antibody generation; ss.
XX
OS   Homo sapiens.
XX
PN   WO9960139-A1.
XX
PD   25-NOV-1999.
XX
PF   29-OCT-1998;   98WO-US022994.
XX
PR   15-MAY-1998;   98US-00079970.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffitt MA,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2000-053300/04.
DR   P-PSDB; AAY55011.
XX
PT   New DNA expression construct for production of enzymatically active
PT   recombinant human beta-tryptase.
XX
PS   Disclosure; Page 43-44; 50pp; English.
XX
CC   This sequence encodes the human beta-tryptase. The invention relates to a
CC   DNA expression construct comprising (5' to 3') a promoter linked to a
CC   signal sequence which is linked to a sequence encoding human beta-
CC   tryptase. The DNA construct is useful for transforming host cells to
CC   express, post translationally process and secrete enzymatically active
CC   human tryptase. The method is useful for the production of large amounts
CC   of tryptase with defined specifications. The transformant is useful for
CC   pharmacological studies, combinatorial library screens and X ray
CC   crystallographic studies. The tryptase produced allows for the
CC   development of tryptase agonists and/or antagonists, is useful as an
CC   antigen to generate antihuman tryptase antibodies in various animals, can
CC   be used in screening for compounds which act as tryptase inhibitors,
CC   antagonists, agonists etc. and to assay for the presence of tryptase in
CC   biological or other solutions. Tryptase inhibitors, antagonists, agonists
CC   etc. may be useful as therapeutics. The tryptase does not require any
CC   post-expression or post-purification modifications or manipulations to
CC   initiate tryptase activity and it has enzymatic activity which compares
CC   favourably with cadaveric tryptase. The availability of enzymatically
CC   active tryptase facilitates the large scale screening of combinatorial
CC   libraries for specific tryptase inhibitors as potential therapeutics and
CC   advances the understanding of the biological significance of tryptase in
CC   mast cell mediated diseases. The tryptase can be used to detect low
CC   levels of tryptase
XX
SQ   Sequence 771 BP; 149 A; 256 C; 242 G; 124 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               7.47e-121       Length:        771
Score:                   1387.00         Matches:       248
Percent Similarity:      99.60%          Conservative:  0
Best Local Similarity:   99.60%          Mismatches:    1
Query Match:             99.57%          Indels:        0
DB:                      3               Gaps:          0

US-09-598-982-21 (1-249) x AAZ40175 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 4
AAS20765
ID   AAS20765 standard; DNA; 771 BP.
XX
AC   AAS20765;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding recombinant human beta-II tryptase.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; gene; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             7. .756
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Recombinant beta-II tryptase"
FT                   /note= "This sequence lacks a start codon"
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12007.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 43; Page 75-77; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   The present sequence encodes for recombinant human beta-II tryptase
XX
SQ   Sequence 771 BP; 149 A; 256 C; 242 G; 124 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              7.47e-121      Length:          771
Score:                  1387.00        Matches:         248
Percent Similarity:     99.60%         Conservative:    0
Best Local Similarity:  99.60%         Mismatches:      1
Query Match:            99.57%         Indels:          0
DB:                     6              Gaps:            0

US-09-598-982-21 (1-249) x AAS20765 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546
```

```
Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 5
AAS20778
ID   AAS20778 standard; DNA; 771 BP.
XX
AC   AAS20778;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant S194A #2.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12012.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 91-92; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 U; 0 Other;
```

```
  Alignment Scores:
  Pred. No.:              1.43e-120     Length:        771
  Score:                  1384.00       Matches:       247
  Percent Similarity:     99.60%        Conservative:  1
  Best Local Similarity:  99.20%        Mismatches:    1
  Query Match:            99.35%        Indels:        0
  DB:                     6             Gaps:          0

US-09-598-982-21 (1-249) x AAS20778 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCCGGCGGA 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCACTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 6
AAS20777
ID   AAS20777 standard; DNA; 771 BP.
XX
AC   AAS20777;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant S194A #1.
XX
```

```
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12011.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 88-90; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 150 A; 255 C; 241 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.43e-120     Length:         771
Score:                  1384.00       Matches:        247
Percent Similarity:     99.60%        Conservative:   1
Best Local Similarity:  99.20%        Mismatches:     1
Query Match:            99.35%        Indels:         0
DB:                     6             Gaps:           0

US-09-598-982-21 (1-249) x AAS20777 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246
```

```
Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCCGGCGGA 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCTCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 7
AAS20763
ID   AAS20763 standard; DNA; 771 BP.
XX
AC   AAS20763;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-I tryptase.
XX
KW   Human; proteolytic tryptase; protease; beta-I tryptase; enzyme; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             7..756
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Beta-I tryptase"
FT                   /note= "This sequence lacks a start codon"
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12006.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
```

```
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Disclosure; Page 72-74; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   The present sequence encodes for human beta-I tryptase
XX
SQ   Sequence 771 BP; 149 A; 257 C; 241 G; 124 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               2.2e-120      Length:        771
Score:                   1382.00       Matches:       247
Percent Similarity:      99.20%        Conservative:  0
Best Local Similarity:   99.20%        Mismatches:    2
Query Match:             99.21%        Indels:        0
DB:                      6             Gaps:          0

US-09-598-982-21 (1-249) x AAS20763 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606
```

```
Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 8
AAS20786
ID   AAS20786 standard; DNA; 771 BP.
XX
AC   AAS20786;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant S194A #6.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12020.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 113-114; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 152 A; 256 C; 240 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:               4.21e-120     Length:        771
Score:                   1379.00       Matches:       246
```

```
Percent Similarity:       99.20%          Conservative:  1
Best Local Similarity:    98.80%          Mismatches:    2
Query Match:              98.99%          Indels:        0
DB:                       6               Gaps:          0

US-09-598-982-21 (1-249) x AAS20786 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCCGGCGGA 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCACTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 9
AAS20785
ID   AAS20785 standard; DNA; 771 BP.
XX
AC   AAS20785;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant S194A #5.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12019.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 110-111; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 150 A; 256 C; 240 G; 125 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.21e-120       Length:         771
Score:                  1379.00         Matches:        246
Percent Similarity:     99.20%          Conservative:   1
Best Local Similarity:  98.80%          Mismatches:     2
Query Match:            98.99%          Indels:         0
DB:                     6               Gaps:           0

US-09-598-982-21 (1-249) x AAS20785 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAG 306
```

```
Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              |||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCCGGCGGA 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCTCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 10
AAS20776
ID   AAS20776 standard; DNA; 771 BP.
XX
AC   AAS20776;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant D91A #1.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12010.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 86-87; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 149 A; 256 C; 242 G; 124 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.21e-120       Length:         771
Score:                  1379.00         Matches:        247
Percent Similarity:     99.20%          Conservative:   0
Best Local Similarity:  99.20%          Mismatches:     2
Query Match:            98.99%          Indels:         0
DB:                     6               Gaps:           0

US-09-598-982-21 (1-249) x AAS20776 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726
```

```
Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 11
AAS20779
ID   AAS20779 standard; DNA; 735 BP.
XX
AC   AAS20779;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant H44A #2.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12013.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Example 1c; Page 94-95; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 735 BP; 139 A; 245 C; 231 G; 120 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.17e-119       Length:        735
Score:                  1374.00         Matches:       245
Percent Similarity:     100.00%         Conservative:  0
Best Local Similarity:  100.00%         Mismatches:    0
Query Match:            98.64%          Indels:        0
DB:                     6               Gaps:          0

US-09-598-982-21 (1-249) x AAS20779 (1-735)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        721 GTCCCCAAAAAGCCG 735

RESULT 12
AAS20784
ID   AAS20784 standard; DNA; 771 BP.
XX
AC   AAS20784;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant D91A #3.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
```

```
PR    21-JUN-2000; 2000US-00598982.
XX
PA    (PROM-) PROMEGA CORP.
XX
PI    Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR    WPI; 2002-114578/15.
DR    P-PSDB; AAU12018.
XX
PT    DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT    comprises DNA sequence encoding proteolytic tryptase having an active
PT    site mutation.
XX
PS    Claim 7; Page 107-109; 126pp; English.
XX
CC    The present invention relates to recombinant human proteolytic tryptases,
CC    active site mutants of these tryptases and the methods for producing
CC    these. The method involves the production of a DNA expression construct
CC    comprising a promoter operably linked to a secretion signal sequence
CC    which is operably linked to a DNA sequence encoding a proteolytic
CC    tryptase with an active site mutation (the construct drives expression of
CC    a mature proteolytic tryptase that lacks enzymatic activity due to the
CC    active site mutation, in hosts transformed to contain the construct). The
CC    method is useful for producing enzymatically-active beta-II tryptase. The
CC    active site mutants of proteolytic tryptase provide a tool to investigate
CC    the structural and functional properties of the protease and its
CC    enzymatic activity, and for modelling studies. The enzymatically-active,
CC    recombinant proteolytic tryptase produced are useful as an antigen to
CC    generate anti-human tryptase antibodies and in drug screening for
CC    compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC    AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC    site mutants
XX
SQ    Sequence 771 BP; 149 A; 257 C; 241 G; 124 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              1.24e-119      Length:          771
Score:                  1374.00        Matches:         246
Percent Similarity:     98.80%         Conservative:    0
Best Local Similarity:  98.80%         Mismatches:      3
Query Match:            98.64%         Indels:          0
DB:                     6              Gaps:            0

US-09-598-982-21 (1-249) x AAS20784 (1-771)

Qy          1 LeuGluLysArgIleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnVal 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          7 CTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTG 66

Qy         21 SerLeuArgValHisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisPro 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         67 AGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCC 126

Qy         41 GlnTrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeu 60
              |||||||||||||||||||||   ||||||||||||||||||||||||||||||||||||
Db        127 CAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTC 186

Qy         61 ArgValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArg 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        187 AGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGG 246

Qy         81 IleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGlu 100
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        247 ATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTGCTGGAG 306

Qy        101 LeuGluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSer 120
              |||||||||||||||   ||||||||||||||||||||||||||||||||||||||||||
Db        307 CTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCA 366

Qy        121 GluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        367 GAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGAT 426

Qy        141 GluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHis 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        427 GAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCAC 486

Qy        161 IleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        487 ATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGT 546

Qy        181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        547 GACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGG 606

Qy        201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        607 CCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAG 666

Qy        221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        667 GGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGG 726

Qy        241 IleHisHisTyrValProLysLysPro 249
              |||||||||||||||||||||||||||
Db        727 ATCCACCACTATGTCCCCAAAAAGCCG 753

RESULT 13
AAS20787
ID    AAS20787 standard; DNA; 735 BP.
XX
AC    AAS20787;
XX
DT    09-APR-2002  (first entry)
XX
DE    DNA encoding human beta-II tryptase active site mutant H44A #4.
XX
KW    Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW    enzyme; mutant; ds.
XX
OS    Homo sapiens.
OS    Synthetic.
XX
PN    WO200198470-A2.
XX
PD    27-DEC-2001.
XX
PF    20-JUN-2001; 2001WO-US019681.
XX
PR    21-JUN-2000; 2000US-00598982.
XX
PA    (PROM-) PROMEGA CORP.
XX
PI    Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR    WPI; 2002-114578/15.
DR    P-PSDB; AAU12021.
XX
PT    DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT    comprises DNA sequence encoding proteolytic tryptase having an active
PT    site mutation.
XX
PS    Example 1c; Page 115-117; 126pp; English.
XX
CC    The present invention relates to recombinant human proteolytic tryptases,
CC    active site mutants of these tryptases and the methods for producing
CC    these. The method involves the production of a DNA expression construct
CC    comprising a promoter operably linked to a secretion signal sequence
CC    which is operably linked to a DNA sequence encoding a proteolytic
CC    tryptase with an active site mutation (the construct drives expression of
CC    a mature proteolytic tryptase that lacks enzymatic activity due to the
CC    active site mutation, in hosts transformed to contain the construct). The
CC    method is useful for producing enzymatically-active beta-II tryptase. The
CC    active site mutants of proteolytic tryptase provide a tool to investigate
CC    the structural and functional properties of the protease and its
CC    enzymatic activity, and for modelling studies. The enzymatically-active,
```

```
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 735 BP; 139 A; 246 C; 230 G; 120 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              3.44e-119      Length:        735
Score:                  1369.00        Matches:       244
Percent Similarity:     99.59%         Conservative:  0
Best Local Similarity:  99.59%         Mismatches:    1
Query Match:            98.28%         Indels:        0
DB:                     6              Gaps:          0

US-09-598-982-21 (1-249) x AAS20787 (1-735)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        721 GTCCCCAAAAAGCCG 735

RESULT 14
AAZ40172
ID   AAZ40172 standard; DNA; 735 BP.
```

```
XX
AC   AAZ40172;
XX
DT   18-FEB-2000  (first entry)
XX
DE   Human beta-tryptase coding sequence.
XX
KW   Beta-tryptase; human; DNA expression construct; protein production;
KW   combinatorial library screening; X ray crystallography; antigen;
KW   antibody generation; ss.
XX
OS   Homo sapiens.
XX
PN   WO9960139-A1.
XX
PD   25-NOV-1999.
XX
PF   29-OCT-1998;   98WO-US022994.
XX
PR   15-MAY-1998;   98US-00079970.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffitt MA,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2000-053300/04.
DR   P-PSDB; AAY55010.
XX
PT   New DNA expression construct for production of enzymatically active
PT   recombinant human beta-tryptase.
XX
PS   Claim 4; Page 40-41; 50pp; English.
XX
CC   This sequence encodes the human beta-tryptase. The invention relates to a
CC   DNA expression construct comprising (5' to 3') a promoter linked to a
CC   signal sequence which is linked to a sequence encoding human beta-
CC   tryptase. The DNA construct is useful for transforming host cells to
CC   express, post translationally process and secrete enzymatically active
CC   human tryptase. The method is useful for the production of large amounts
CC   of tryptase with defined specifications. The transformant is useful for
CC   pharmacological studies, combinatorial library screens and X ray
CC   crystallographic studies. The tryptase produced allows for the
CC   development of tryptase agonists and/or antagonists, is useful as an
CC   antigen to generate antihuman tryptase antibodies in various animals, can
CC   be used in screening for compounds which act as tryptase inhibitors,
CC   antagonists, agonists etc. and to assay for the presence of tryptase in
CC   biological or other solutions. Tryptase inhibitors, antagonists, agonists
CC   etc. may be useful as therapeutics. The tryptase does not require any
CC   post-expression or post-purification modifications or manipulations to
CC   initiate tryptase activity and it has enzymatic activity which compares
CC   favourably with cadaveric tryptase. The availability of enzymatically
CC   active tryptase facilitates the large scale screening of combinatorial
CC   libraries for specific tryptase inhibitors as potential therapeutics and
CC   advances the understanding of the biological significance of tryptase in
CC   mast cell mediated diseases. The tryptase can be used to detect low
CC   levels of tryptase
XX
SQ   Sequence 735 BP; 141 A; 245 C; 229 G; 120 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.27e-119      Length:        735
Score:                  1368.00        Matches:       244
Percent Similarity:     99.59%         Conservative:  0
Best Local Similarity:  99.59%         Mismatches:    1
Query Match:            98.21%         Indels:        0
DB:                     3              Gaps:          0

US-09-598-982-21 (1-249) x AAZ40172 (1-735)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60
```

```
Qy          25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy          45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
               |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy          65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy          85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy         105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy         125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy         145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy         165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy         185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy         205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy         225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy         245 ValProLysLysPro 249
               |||||||||||||||
Db         721 GTCCCCAAAAAGCCG 735

RESULT 15
AAS20766
ID   AAS20766 standard; DNA; 735 BP.
XX
AC   AAS20766;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding recombinant human mature beta-II tryptase.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; gene; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             1. .735
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Recombinant mature beta-II tryptase"
FT                   /note= "This sequence lacks both start and stop codons"
XX
PN   WO200198470-A2.
XX
```

```
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12008.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Disclosure; Page 78-79; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   The present sequence encodes for recombinant human mature beta-II
CC   tryptase
XX
SQ   Sequence 735 BP; 141 A; 245 C; 229 G; 120 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:              4.27e-119       Length:         735
Score:                  1368.00         Matches:        244
Percent Similarity:     99.59%          Conservative:   0
Best Local Similarity:  99.59%          Mismatches:     1
Query Match:            98.21%          Indels:         0
DB:                     6               Gaps:           0

US-09-598-982-21 (1-249) x AAS20766 (1-735)

Qy           5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy          25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy          45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
               |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy          65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy          85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy         105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360
```

```
Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        721 GTCCCCAAAAAGCCG 735

Search completed: July 22, 2004, 16:02:43
Job time : 479 secs
```

```
                 GenCore version 5.1.6
        Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         July 22, 2004, 17:23:03 ; Search time 520 Seconds
                                           (without alignments)
                                           2340.743 Million cell updates/sec

Title:          US-09-598-982-21
Perfect score:  1393
Sequence:       1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       3216467 seqs, 2444149694 residues

Total number of hits satisfying chosen parameters:      6432934

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_20072004_064412_6622/app_query.fasta_1.391
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09598982_@CGN_1_1_912_@runat_20072004_064412_6622
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
                14:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                15:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                16:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
                17:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                18:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                19:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                  %
Result           Query
  No.    Score   Match Length DB  ID                     Description
```

```
-------------------------------------------------------------------------------
     1    1368    98.2    1143  16  US-10-352-684A-47          Sequence 47, Appl
     2    1368    98.2    1145  17  US-10-287-226-93           Sequence 93, Appl
     3    1363    97.8     828  17  US-10-287-226-91           Sequence 91, Appl
     4    1344    96.5    1081  9  US-09-954-456-2126         Sequence 2126, Ap
     5    1344    96.5    1081  10  US-09-960-706-680          Sequence 680, App
     6    1344    96.5    1081  10  US-09-873-319-427          Sequence 427, App
     7    1344    96.5    1081  10  US-09-873-367C-155         Sequence 155, App
     8    1344    96.5    1081  10  US-09-873-367C-714         Sequence 714, App
     9    1329    95.4    2662  17  US-10-275-505-27           Sequence 27, Appl
    10    1323    95.0     828  17  US-10-287-226-89           Sequence 89, Appl
    11    1312    94.2     858  17  US-10-287-226-87           Sequence 87, Appl
    12    1274    91.5    1154  16  US-10-352-684A-45          Sequence 45, Appl
    13    1274    91.5    1158  13  US-10-116-802-240          Sequence 240, App
    14    1054    75.7    1187  17  US-10-275-505-22           Sequence 22, Appl
    15   664.5    47.7     846  10  US-09-813-432-11           Sequence 11, Appl
    16   664.5    47.7     846  13  US-10-246-583-11           Sequence 11, Appl
    17   664.5    47.7     846  16  US-10-174-364-11           Sequence 11, Appl
    18   664.5    47.7     846  17  US-10-689-832-11           Sequence 11, Appl
    19     634    45.5     873  15  US-10-117-323-2            Sequence 2, Appli
    20   605.5    43.5    1122  9  US-09-900-754-1            Sequence 1, Appli
    21   599.5    43.0     948  13  US-10-150-813-15           Sequence 15, Appl
    22   599.5    43.0     948  13  US-10-139-854-15           Sequence 15, Appl
    23   599.5    43.0     948  15  US-10-131-409-15           Sequence 15, Appl
    24   599.5    43.0     948  16  US-10-150-811-15           Sequence 15, Appl
    25     584    41.9     978  15  US-10-311-955-1            Sequence 1, Appli
    26   583.5    41.9    1958  17  US-10-311-035-29           Sequence 29, Appl
    27   569.5    40.9     867  9  US-09-888-615-56           Sequence 56, Appl
    28   569.5    40.9     895  15  US-10-117-323-34           Sequence 34, Appl
    29   569.5    40.9    1218  16  US-10-274-639-36           Sequence 36, Appl
    30   569.5    40.9    1218  17  US-10-333-574-36           Sequence 36, Appl
    31     568    40.8     858  13  US-10-150-813-21           Sequence 21, Appl
    32     568    40.8     858  13  US-10-139-854-21           Sequence 21, Appl
    33     568    40.8     858  15  US-10-131-409-21           Sequence 21, Appl
    34     568    40.8     858  16  US-10-150-811-21           Sequence 21, Appl
    35   563.5    40.5     780  15  US-10-266-035-3            Sequence 3, Appli
    36   563.5    40.5     783  17  US-10-391-364-22           Sequence 22, Appl
    37   563.5    40.5     980  15  US-10-266-035-1            Sequence 1, Appli
    38   563.5    40.5     980  15  US-10-040-647-30           Sequence 30, Appl
    39   563.5    40.5     980  17  US-10-391-364-20           Sequence 20, Appl
    40   563.5    40.5    1110  14  US-10-040-655-1            Sequence 1, Appli
    41   563.5    40.5    1110  14  US-10-041-006A-1           Sequence 1, Appli
    42   563.5    40.5    1129  13  US-10-147-493-221          Sequence 221, App
    43   563.5    40.5    1129  13  US-10-145-127-221          Sequence 221, App
    44   563.5    40.5    1129  13  US-10-160-503-221          Sequence 221, App
    45   563.5    40.5    1129  13  US-10-143-118-221          Sequence 221, App


                                  ALIGNMENTS

RESULT 1
US-10-352-684A-47
; Sequence 47, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
```

```
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
;   LENGTH: 1143
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (6)...(833)
US-10-352-684A-47

Alignment Scores:
Pred. No.:                2.3e-158      Length:          1143
Score:                    1368.00       Matches:         244
Percent Similarity:       99.59%        Conservative:    0
Best Local Similarity:    99.59%        Mismatches:      1
Query Match:              98.21%        Indels:          0
DB:                       16            Gaps:            0

US-09-598-982-21 (1-249) x US-10-352-684A-47 (1-1143)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 155

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 215

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 275

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 335

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 395

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        396 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 455

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 515

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        516 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 575

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 635
```

```
Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        636 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 695

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        696 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 755

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        756 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 815

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        816 GTCCCCAAAAAGCCG 830

RESULT 2
US-10-287-226-93
; Sequence 93, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
```

```
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 93
;   LENGTH: 1145
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (8)..(832)
US-10-287-226-93

Alignment Scores:
Pred. No.:              2.31e-158       Length:         1145
Score:                  1368.00         Matches:        244
Percent Similarity:     99.59%          Conservative:   0
Best Local Similarity:  99.59%          Mismatches:     1
Query Match:            98.21%          Indels:         0
DB:                     17              Gaps:           0

US-09-598-982-21 (1-249) x US-10-287-226-93 (1-1145)

Qy            5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           98 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 157

Qy           25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          158 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 217

Qy           45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
                |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db          218 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 277

Qy           65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          278 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 337

Qy           85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          338 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 397

Qy          105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          398 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 457

Qy          125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          458 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 517

Qy          145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          518 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 577

Qy          165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          578 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 637

Qy          185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          638 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 697

Qy          205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          698 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 757

Qy          225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          758 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 817

Qy          245 ValProLysLysPro 249
                |||||||||||||||
Db          818 GTCCCCAAAAAGCCG 832

RESULT 3
US-10-287-226-91
; Sequence 91, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
```

```
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 91
;   LENGTH: 828
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(825)
US-10-287-226-91

Alignment Scores:
Pred. No.:              6.14e-158     Length:        828
Score:                  1363.00       Matches:       243
Percent Similarity:     99.18%        Conservative:  0
Best Local Similarity:  99.18%        Mismatches:    2
Query Match:            97.85%        Indels:        0
DB:                     17            Gaps:          0

US-09-598-982-21 (1-249) x US-10-287-226-91 (1-828)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 150

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 210

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        211 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 270

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 330

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        331 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 390

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        391 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 450

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        451 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 510

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 570

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 630

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 690

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 750

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        751 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 810
```

```
Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        811 GTCCCCAAAAAGCCG 825

RESULT 4
US-09-954-456-2126
; Sequence 2126, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cance
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2126
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-2126

Alignment Scores:
Pred. No.:              1.94e-155     Length:        1081
Score:                  1344.00       Matches:       241
Percent Similarity:     98.37%        Conservative:  0
Best Local Similarity:  98.37%        Mismatches:    4
Query Match:            96.48%        Indels:        0
DB:                     9             Gaps:          0

US-09-598-982-21 (1-249) x US-09-954-456-2126 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367
```

```
Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 5
US-09-960-706-680
; Sequence 680, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
;  TITLE OF INVENTION:  Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-01US
;  CURRENT APPLICATION NUMBER: US/09/960,706
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 60/223,323
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: 09/873,319
;  PRIOR FILING DATE: 2001-06-05
;  NUMBER OF SEQ ID NOS: 1124
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134280A1 M33493
US-09-960-706-680

 Alignment Scores:
 Pred. No.:               1.94e-155     Length:        1081
 Score:                   1344.00       Matches:       241
 Percent Similarity:      98.37%        Conservative:  0
 Best Local Similarity:   98.37%        Mismatches:    4
 Query Match:             96.48%        Indels:        0
 DB:                      10            Gaps:          0

 US-09-598-982-21 (1-249) x US-09-960-706-680 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 6
US-09-873-319-427
; Sequence 427, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  APPLICANT: Kulkarni, Prakash
;  APPLICANT: Getzenberg, Robert H.
;  APPLICANT: Waga, Iwao
;  APPLICANT: Yamamoto, Jun
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
;  TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-US
;  CURRENT APPLICATION NUMBER: US/09/873,319A
;  CURRENT FILING DATE: 2001-06-05
;  EARLIER APPLICATION NUMBER: US 60/223,323
;  EARLIER FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 755
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 427
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134324A1 M33493
US-09-873-319-427
```

```
Alignment Scores:
Pred. No.:                 1.94e-155     Length:         1081
Score:                     1344.00       Matches:        241
Percent Similarity:        98.37%        Conservative:   0
Best Local Similarity:     98.37%        Mismatches:     4
Query Match:               96.48%        Indels:         0
DB:                        10            Gaps:           0

US-09-598-982-21 (1-249) x US-09-873-319-427 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||    |||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 7
US-09-873-367C-155
; Sequence 155, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-155

Alignment Scores:
Pred. No.:                 1.94e-155     Length:         1081
Score:                     1344.00       Matches:        241
Percent Similarity:        98.37%        Conservative:   0
Best Local Similarity:     98.37%        Mismatches:     4
Query Match:               96.48%        Indels:         0
DB:                        10            Gaps:           0

US-09-598-982-21 (1-249) x US-09-873-367C-155 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||    |||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 8
US-09-873-367C-714
; Sequence 714, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 714
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-714

Alignment Scores:
Pred. No.:              1.94e-155     Length:          1081
Score:                  1344.00       Matches:         241
Percent Similarity:     98.37%        Conservative:    0
Best Local Similarity:  98.37%        Mismatches:      4
Query Match:            96.48%        Indels:          0
DB:                     10            Gaps:            0

US-09-598-982-21 (1-249) x US-09-873-367C-714 (1-1081)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         68 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 127

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        128 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 187

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        188 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 247

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 307

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 367

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        368 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 427

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 487

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        488 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 547

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        548 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 607

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        608 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 667

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 727

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        728 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 787

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        788 GTCCCCAAAAAGCCG 802

RESULT 9
US-10-275-505-27
; Sequence 27, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  DELEGEANE, Angelo M.; LAL, Preeti G.
;  APPLICANT:  HAFALIA, April J.A.; PATTERSON, Chandra
;  APPLICANT:  WALIA, Narinder K.; KEARNEY, Liam
;  APPLICANT:  TRIBOULEY, Catherine M.; KHAN, Farrah A.
;  APPLICANT:  YAO, Monique G.; BAUGHN, Mariah R.
;  APPLICANT:  AZIMZAI, Yalda; ELLIOTT, Vicki S.
;  APPLICANT:  NGUYEN, Danniel B.; GANDHI, Ameena R.
;  APPLICANT:  YANG, Junming; HERNANDEZ, Roberto
;  APPLICANT:  POLICKY, Jennifer L.; LU, Dyung Aina M.
;  APPLICANT:  REDDY, Roopa M.; YUE, Henry
;  APPLICANT:  TANG, Y. Tom
;  TITLE OF INVENTION: PROTEASES
;  FILE REFERENCE: PI-0085 USN
;  CURRENT APPLICATION NUMBER: US/10/275,505
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: PCT/US01/14651
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: 60/209,402
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: 60/207,477
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/205,803
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/203,566
;  PRIOR FILING DATE: 2000-05-11
;  PRIOR APPLICATION NUMBER: 60/202,082
;  PRIOR FILING DATE: 2000-05-04
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PERL Program
; SEQ ID NO 27
;   LENGTH: 2662
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7472460CB1
```

```
US-10-275-505-27

Alignment Scores:
Pred. No.:              4.65e-153      Length:        2662
Score:                  1329.00        Matches:       238
Percent Similarity:     97.14%         Conservative:  0
Best Local Similarity:  97.14%         Mismatches:    7
Query Match:            95.41%         Indels:        0
DB:                     17             Gaps:          0

US-09-598-982-21 (1-249) x US-10-275-505-27 (1-2662)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        701 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 760

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        761 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 820

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        821 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 880

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        881 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 940

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        941 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 1000

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1001 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 1060

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1061 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 1120

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||   |||||||||||||||||||||||||||||||||||||||
Db       1121 CCGCCATTTCCTCTGAAGCATGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 1180

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1181 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 1240

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1241 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 1300

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||   |||||||||||||||||||||
Db       1301 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGATGGGGAGAGGGCTGTGCCCAG 1360

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1361 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 1420

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db       1421 GTCCCCAAAAAGCCG 1435

RESULT 10
US-10-287-226-89
; Sequence 89, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
```

```
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 89
;   LENGTH: 828
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(825)
US-10-287-226-89

Alignment Scores:
Pred. No.:              5.21e-153      Length:        828
Score:                  1323.00        Matches:       234
Percent Similarity:     97.14%         Conservative:  4
```

```
Best Local Similarity:  95.51%          Mismatches:    7
Query Match:            94.97%          Indels:        0
DB:                     17              Gaps:          0

US-09-598-982-21 (1-249) x US-10-287-226-89 (1-828)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         91 ATTGTTGGGGGGCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 150

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 CGCGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTA 210

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||   ||||||:::||||||||||||||||||||||||||||||
Db        211 ACCGCGGCGCACTGCGTGGAACCGGACATCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 270

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        271 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 330

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db        331 CCACAGTTCTACATCATCCAGACCGGGGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCC 390

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::|||||||||:::|||||||||||||||||||||||||||||||||||||||
Db        391 GTGAACATCTCCAGCCACATCCACACGGTCACGCTGCCCCCTGCCTCGGAGACCTTCCCC 450

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||||
Db        451 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGACGTGGACAATAATGAGCGCCTCCCA 510

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        511 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 570

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        571 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 630

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        631 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 690

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        691 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 750

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        751 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 810

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        811 GTCCCCAAAAAGCCG 825

RESULT 11
US-10-287-226-87
; Sequence 87, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 87
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (8)..(844)
US-10-287-226-87

Alignment Scores:
Pred. No.:                1.24e-151       Length:          858
Score:                    1312.00         Matches:         238
Percent Similarity:       96.75%          Conservative:    0
Best Local Similarity:    96.75%          Mismatches:      2
Query Match:              94.19%          Indels:          6
DB:                       17              Gaps:            2

US-09-598-982-21 (1-249) x US-10-287-226-87 (1-858)
```

```
Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 181

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 241

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 301

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 361

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 421

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 481

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              |||||||||||||||||||||||||||||||||||||||               ||||||
Db        482 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTG---------------CTCCCA 526

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        527 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 586

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        587 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 646

Qy        185 CysAlaGlyAsnThrArgArgAspSerCys---GlnGlyAspSerGlyGlyProLeuVal 203
              ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db        647 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGCAGGGCGACTCCGGAGGGCCCCTGGTG 706

Qy        204 CysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAla 223
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        707 TGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCC 766

Qy        224 GlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHis 243
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        767 CAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCAC 826

Qy        244 TyrValProLysLysPro 249
              ||||||||||||||||||
Db        827 TATGTCCCCAAAAAGCCG 844

RESULT 12
US-10-352-684A-45
; Sequence 45, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 1154
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (18)...(845)
US-10-352-684A-45

Alignment Scores:
Pred. No.:               8.93e-147       Length:         1154
Score:                   1274.00         Matches:        226
Percent Similarity:      94.69%          Conservative:   6
Best Local Similarity:   92.24%          Mismatches:     13
Query Match:             91.46%          Indels:         0
DB:                      16              Gaps:           0

US-09-598-982-21 (1-249) x US-10-352-684A-45 (1-1154)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 167

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 227

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   |||:::||||||||||||||||||||||||   |||||||||||||||
Db        228 ACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTG 287

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGCAC 347

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db        348 CCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCC 407

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::||||||   ||||||||||||   |||||||||||||||||||||||||||
Db        408 GTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCC 467

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        468 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCA 527

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 CCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 587

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
```

```
Db        588 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTG 647

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||::::::||||||||||||:::|||||||||||||||||||||||||||
Db        648 TGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGGCCCCTGGTGTGC 707

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        708 AAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCCAG 767

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        768 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 827

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        828 GTCCCCAAAAAGCCG 842

RESULT 13
US-10-116-802-240
; Sequence 240, Application US/10116802
; Publication No. US20030065157A1
; GENERAL INFORMATION:
;  APPLICANT: Amy Lasek
;  TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
;  FILE REFERENCE: PA-0045 US
;  CURRENT APPLICATION NUMBER: US/10/116,802
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 60/281,593
;  PRIOR FILING DATE: 2001-04-04
;  NUMBER OF SEQ ID NOS: 519
;  SOFTWARE: PERL Program
; SEQ ID NO 240
;   LENGTH: 1158
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 378633.40
US-10-116-802-240

Alignment Scores:
Pred. No.:              8.98e-147      Length:         1158
Score:                  1274.00        Matches:        226
Percent Similarity:     94.69%         Conservative:   6
Best Local Similarity:  92.24%         Mismatches:     13
Query Match:            91.46%         Indels:         0
DB:                     13             Gaps:           0

US-09-598-982-21 (1-249) x US-10-116-802-240 (1-1158)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        108 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 167

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        168 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 227

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   |||:::||||||||||||||||||||||||   |||||||||||||||
Db        228 ACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTG 287

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        288 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGCAC 347

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db        348 CCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCC 407

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::||||||   ||||||||||||   |||||||||||||||||||||||||||
```

```
Db        408 GTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCC 467

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        468 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCA 527

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        528 CCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 587

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        588 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTG 647

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||::::::||||||||||||:::|||||||||||||||||||||||||||
Db        648 TGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGGCCCCTGGTGTGC 707

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db        708 AAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCCAG 767

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        768 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 827

Qy        245 ValProLysLysPro 249
              |||||||||||||||
Db        828 GTCCCCAAAAAGCCG 842

RESULT 14
US-10-275-505-22
; Sequence 22, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  DELEGEANE, Angelo M.; LAL, Preeti G.
;  APPLICANT:  HAFALIA, April J.A.; PATTERSON, Chandra
;  APPLICANT:  WALIA, Narinder K.; KEARNEY, Liam
;  APPLICANT:  TRIBOULEY, Catherine M.; KHAN, Farrah A.
;  APPLICANT:  YAO, Monique G.; BAUGHN, Mariah R.
;  APPLICANT:  AZIMZAI, Yalda; ELLIOTT, Vicki S.
;  APPLICANT:  NGUYEN, Danniel B.; GANDHI, Ameena R.
;  APPLICANT:  YANG, Junming; HERNANDEZ, Roberto
;  APPLICANT:  POLICKY, Jennifer L.; LU, Dyung Aina M.
;  APPLICANT:  REDDY, Roopa M.; YUE, Henry
;  APPLICANT:  TANG, Y. Tom
;  TITLE OF INVENTION: PROTEASES
;  FILE REFERENCE: PI-0085 USN
;  CURRENT APPLICATION NUMBER: US/10/275,505
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: PCT/US01/14651
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: 60/209,402
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: 60/207,477
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/205,803
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/203,566
;  PRIOR FILING DATE: 2000-05-11
;  PRIOR APPLICATION NUMBER: 60/202,082
;  PRIOR FILING DATE: 2000-05-04
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PERL Program
; SEQ ID NO 22
;   LENGTH: 1187
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 3897384CB1
US-10-275-505-22
```

```
Alignment Scores:
Pred. No.:              1.19e-119      Length:        1187
Score:                  1054.00        Matches:       186
Percent Similarity:     83.54%         Conservative:  17
Best Local Similarity:  76.54%         Mismatches:    40
Query Match:            75.66%         Indels:        0
DB:                     17             Gaps:          0

US-09-598-982-21 (1-249) x US-10-275-505-22 (1-1187)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              |||||||||||||||||||||      :::|||||||||||||||||||||||||||
Db        263 ATTGTGGGGGACAGGAGGCACATGGGAACAAGTGGCCCTGGCAGGTGAGCCTGCGTGCC 322

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              :::      |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        323 AATGACACCTACTGGATGCATTTCTGCGGTGGCTCCCTCATCCACCCACAGTGGGTGCTC 382

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||   |||         :::||||||||||||
Db        383 ACTGCGGCACACTGTGTGGGACCGGATGTTGCTGACCCCAACAAGGTCAGAGTACAGCTC 442

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              |||:::|||:::|||||||||   |||   |||:::   ||||||:::||||||   |||
Db        443 CGTAAGCAGTACCTCTATTACCATGACCACCTGATGACTGTGAGCCAGATCATCACACAC 502

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              |||   ||||||      |||   |||||||||||||||||||||:::|||      |||
Db        503 CCCGACTTCTACATCGTCCAGGATGGGGCAGACATTGCCCTGCTGAAACTCACAAACCCT 562

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::|||   :::||||||   |||   |||||||||||||||||||||||||||
Db        563 GTGAACATTTCTGACTATGTCCACCCTGTCCCCCTACCTCCTGCCTCAGAGACCTTCCCC 622

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
                 |||      |||||||||||||||||||||::::::||||||         ||||||
Db        623 TCAGGAACGTTGTGCTGGGTGACAGGCTGGGGTAACATCGACAATGGTGTAAACCTGCCG 682

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||:::|||:::|||||||||:::|||||||||:::||||||
Db        683 CCACCATTTCCTTTGAAGGAGGTGCAAGTTCCCATTATAGAAAACCACCTTTGTGACTTG 742

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              |||||||||   |||      |||||||||:::|||   |||||||||||||||||||||
Db        743 AAGTATCACAAAGGTCTCATCACAGGTGACAATGTCCACATTGTCCGAGATGACATGCTG 802

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||         |||||||||||||||||||||||||||||||||||||||
Db        803 TGTGCTGGGAATGAAGGACATGACTCCTGCCAGGGCGACTCCGGAGGACCTCTGGTCTGC 862

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGln 224
              ||||||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db        863 AAGGTAGAAGACACCTGGCTGCAGGCAGGCGTGGTCAGCTGGGGTGAGGGCTGTGCACAG 922

Qy        225 ProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyr 244
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        923 CCCAACAGGCCTGGCATCTACACCCGGGTCACCTATTACTTGGACTGGATCCACCACTAT 982

Qy        245 ValProLys 247
              |||||||||
Db        983 GTCCCCAAG 991

RESULT 15
US-09-813-432-11
; Sequence 11, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier Jr., Raymond J
;  APPLICANT:  Majmuder, Kamud
;  APPLICANT:  Spaderna, Steven K
;  APPLICANT:  Smithson, Glenda
;  APPLICANT:  Mezes, Peter S
;  APPLICANT:  Vernet, Corine A. M.
;  TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
;  FILE REFERENCE: 15966-729
;  CURRENT APPLICATION NUMBER: US/09/813,432
;  CURRENT FILING DATE:  2001-03-20
;  PRIOR APPLICATION NUMBER: 60/190,835
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,768
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,972
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,199
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,947
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,665
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,657
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,984
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,664
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,836
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: 60/193,843
;  PRIOR FILING DATE: 2000-03-31
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 846
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-813-432-11

Alignment Scores:
Pred. No.:              7.46e-72       Length:        846
Score:                  664.50         Matches:       129
Percent Similarity:     65.06%         Conservative:  33
Best Local Similarity:  51.81%         Mismatches:    82
Query Match:            47.70%         Indels:        5
DB:                     10             Gaps:          3

US-09-598-982-21 (1-249) x US-09-813-432-11 (1-846)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||      ||||||      ||||||||||||||||||||||||||||||
Db         96 ATTGTGGGGGGCCACAATGCCCCCCCGGGGAAGTGGCCGTGGCAGGTCAGCCTGAGGGTC 155

Qy         25 ---------HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGln 41
                       |||         |||   |||   |||||||||||||||||||||||||||
Db        156 TACAGCTACCACTGGGCCTCCTGGGCGCACATCTGTGGGGGCTCCCTCATCCACCCCCAG 215

Qy         42 TrpValLeuThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArg 61
              ||||||||||||||||||   |||:::               |||   :::      |||
Db        216 TGGGTGCTGACTGCTGCCCACTGCATTTTCTGGAAGGACACCGACCCGTCCATCTACCGG 275

Qy         62 ValGlnLeuArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIle 81
              :::         :::   :::||||||            ||||||   ||||||||||||
Db        276 ATCCACGCTGGGGACGTGTATCTCTACGGGGGCCGGGGGCTGCTGAACGTCAGCCGGATC 335

Qy         82 IleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeu 101
              ||||||||||||   :::   ||||||   :::|||||||||:::|||||||||:::|||
Db        336 ATCGTCCACCCCAACTATGTCACTGCGGGGCTGGGTGCGGATGTGGCCCTGCTCCAGCTG 395

Qy        102 GluGluProValLysValSerSerHisValHisThrValThrLeuProProAlaSerGlu 121
                    |||:::      :::::::::|||   ||||||   |||   |||   |||
Db        396 GTGAGCCCCATGATCGGAGCCGCTAATGTCAGGACGGTCAAGCTCTCCCCGGTCTCGCTG 455

Qy        122 ThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGlu 141
                       |||         |||||||||||||||||||||   :::         |||
Db        456 GAGCTCACCCCGAAGGACCAGTGCTGGGTGACTGGCTGGGGAGCGATCAGGATGTTCGAG 515
```

```
Qy          142 ArgLeuProProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIle 161
                   ||||||||||||:::   |||:::|||      |||   ::::::||||||   :::
Db          516 TCGCTGCCGCCGCCCTACCGCCTGCAGCAGGCGAGTGTGCAGGTGCTGGAGAACGCCGTC 575

Qy          162 CysAspAlaLysTyr---HisLeuGlyAlaTyrThrGlyAspAspValArgIleValArg 180
                |||:::      |||   :::         :::|||||||||      :::::::::
Db          576 TGTGAGCAGCCCTACCGCAACGCCTCAGGGCACACTGGCGACCGG---CAGCTCATCCTG 632

Qy          181 AspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGly 200
                |||||||||||||||||||||:::      ||||||||||||||||||||||||||||||
Db          633 GATGACATGCTGTGTGCCGGCAGCGAGGGCCGAGACTCCTGTCAGGGTGACTCCGGCGGC 692

Qy          201 ProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGlu 220
                ||||||||||||::::::   |||:::|||         ||||||||||||||||||
Db          693 CCTCTGGTCTGCAGGCTGCGGGGGTCCTGGCGCCTGGTGGGGGTGGTCAGCTGGGGCTAC 752

Qy          221 GlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrp 240
                ||||||         :::   ||||||:::||||||   |||      |||:::   |||
Db          753 GGCTGTACCCTGCGGGACTTTCCCGGCGTCTACACCCACGTCCAGATCTACGTGCTCTGG 812

Qy          241 IleHisHisTyrValProLysLysPro 249
                |||         |||   :::   |||
Db          813 ATCCTGCAGCAAGTCGGGGAGTTGCCC 839

Search completed: July 22, 2004, 19:33:58
Job time : 526 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - protein search, using sw model

Run on:          July 22, 2004, 14:36:26 ; Search time 83 Seconds
                                           (without alignments)
                                           946.555 Million cell updates/sec

Title:           US-09-598-982-21
Perfect score:   1393
Sequence:        1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched:        1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters:      1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       SPTREMBL_25:*
                 1:  sp_archea:*
                 2:  sp_bacteria:*
                 3:  sp_fungi:*
                 4:  sp_human:*
                 5:  sp_invertebrate:*
                 6:  sp_mammal:*
                 7:  sp_mhc:*
                 8:  sp_organelle:*
                 9:  sp_phage:*
                 10:  sp_plant:*
                 11:  sp_rodent:*
                 12:  sp_virus:*
                 13:  sp_vertebrate:*
                 14:  sp_unclassified:*
                 15:  sp_rvirus:*
                 16:  sp_bacteriap:*
                 17:  sp_archeap:*

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.

                                 SUMMARIES
                     %
Result            Query
  No.    Score    Match Length DB  ID                       Description
-----------------------------------------------------------------------------
   1      1342    96.3    275  4  Q96RZ6                   Q96rz6 homo sapien
   2      1286    92.3    275  4  Q86TM8                   Q86tm8 homo sapien
   3      1286    92.3    276  4  Q86UA5                   Q86ua5 homo sapien
   4    1092.5    78.4    275  6  Q7YS62                   Q7ys62 equus cabal
   5      1051    75.4    271  6  Q8HYJ2                   Q8hyj2 bos taurus
   6      1046    75.1    273  11  Q921N4                   Q921n4 mus musculu
   7      1022    73.4    233  4  Q96RZ7                   Q96rz7 homo sapien
   8      1019    73.2    273  6  Q9XSM1                   Q9xsm1 ovis aries
   9      1005    72.1    237  6  Q29464                   Q29464 bos taurus
  10       678    48.7    318  11  Q80UR4                   Q80ur4 mus musculu
  11       651    46.7    277  6  Q8SQ44                   Q8sq44 sus scrofa
  12       640    45.9    279  11  Q99MS4                   Q99ms4 mus musculu
  13     596.5    42.8    328  11  Q8BJR6                   Q8bjr6 mus musculu
  14     589.5    42.3    321  4  Q96RZ8                   Q96rz8 homo sapien
  15     587.5    42.2    277  11  Q80WM7                   Q80wm7 mus musculu
  16       568    40.8    274  11  Q924N9                   Q924n9 mus musculu
```

```
  17     564.5    40.5    311  11  Q80XZ3                   Q80xz3 rattus norv
  18     549.5    39.4    320  13  Q7T0X2                   Q7t0x2 xenopus lae
  19       543    39.0    389  13  Q9PVX7                   Q9pvx7 xenopus lae
  20     538.5    38.7    327  4  Q8N171                   Q8n171 homo sapien
  21       537    38.5    331  11  Q8R1A6                   Q8r1a6 mus musculu
  22       537    38.5    331  11  Q80X17                   Q80x17 mus musculu
  23     536.5    38.5    284  4  Q8NF86                   Q8nf86 homo sapien
  24     536.5    38.5    310  11  Q91XC4                   Q91xc4 mus musculu
  25     534.5    38.4    317  13  Q9DGR3                   Q9dgr3 xenopus lae
  26     531.5    38.2    310  11  Q9QYZ9                   Q9qyz9 mus musculu
  27       530    38.0    339  11  Q99L44                   Q99l44 mus musculu
  28     521.5    37.4    340  11  Q8BJV6                   Q8bjv6 mus musculu
  29       494    35.5    328  11  Q80Z40                   Q80z40 rattus norv
  30     492.5    35.4    297  11  O88781                   O88781 rattus ratt
  31       492    35.3    453  11  Q812A6                   Q812a6 mus musculu
  32       488    35.0    371  11  Q8CJ16                   Q8cj16 rattus norv
  33       488    35.0    445  11  Q8CJ17                   Q8cj17 rattus norv
  34       481    34.5    455  11  Q8CDR0                   Q8cdr0 mus musculu
  35       479    34.4    471  11  Q8CFE0                   Q8cfe0 mus musculu
  36     478.5    34.4    307  11  Q7TML0                   Q7tml0 mus musculu
  37       476    34.2    336  11  Q80YD8                   Q80yd8 mus musculu
  38       473    34.0    558  4  Q86YM4                   Q86ym4 homo sapien
  39       472    33.9    329  13  O42272                   O42272 xenopus lae
  40     469.5    33.7    335  11  Q8VIF2                   Q8vif2 mus musculu
  41       469    33.7    624  6  Q95ME7                   Q95me7 oryctolagus
  42       468    33.6    537  4  Q9BYE1                   Q9bye1 homo sapien
  43       468    33.6    581  4  Q9BYE2                   Q9bye2 homo sapien
  44     463.5    33.3    624  11  Q91Y47                   Q91y47 mus musculu
  45     460.5    33.1    624  11  Q9DAT3                   Q9dat3 mus musculu


                                ALIGNMENTS


RESULT 1
Q96RZ6
ID   Q96RZ6      PRELIMINARY;      PRT;   275 AA.
AC   Q96RZ6;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tryptase I.
GN   TRYPTASEC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21096910; PubMed=11157797;
RA   Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA   Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA   Higgs D.R.;
RT   "Sequence, structure and pathology of the fully annotated terminal 2
RT   Mb of the short arm of human chromosome 16.";
RL   Hum. Mol. Genet. 10:339-352(2001).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AE006466; AAK61271.1; -.
DR   HSSP; P00761; 1AN1.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
```

```
SQ   SEQUENCE   275 AA;  30601 MW;  2F4AAAAE7DDE483E CRC64;

  Query Match             96.3%;  Score 1342;  DB 4;  Length 275;
  Best Local Similarity   98.0%;  Pred. No. 2.4e-123;
  Matches  240;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
                ||||||||||||||||||||| |||||||||||||||||||||| |||||||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVHDPYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||| |
Db        211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVRWGEGCAQPNRPGIYTRVTYYLDWIHXY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275

RESULT 2
Q86TM8
ID   Q86TM8       PRELIMINARY;      PRT;   275 AA.
AC   Q86TM8;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to tryptase, alpha.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain, Lung, and Testis;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC028059; AAH28059.1; -.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
SQ   SEQUENCE   275 AA;  30713 MW;  189C3EBBCB473FA3 CRC64;

  Query Match             92.3%;  Score 1286;  DB 4;  Length 275;
  Best Local Similarity   93.5%;  Pred. No. 7.4e-118;
  Matches  229;  Conservative    3;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| |:|||||||| |||||
Db         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
```

```
Db        151 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 210

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        211 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 270

Qy        245 VPKKP 249
              |||||
Db        271 VPKKP 275

RESULT 3
Q86UA5
ID   Q86UA5       PRELIMINARY;      PRT;   276 AA.
AC   Q86UA5;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to tryptase, alpha (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC051852; AAH51852.1; -.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   276 AA;  30869 MW;  8CDC5EA037955F9C CRC64;

  Query Match             92.3%;  Score 1286;  DB 4;  Length 276;
  Best Local Similarity   93.5%;  Pred. No. 7.4e-118;
  Matches  229;  Conservative    3;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
```

```
               ||||||||||||||||||||   |||||||||||||||||||| |:|||||||| |||||
Db          32 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCLGPDVKDLATLRVQL 91

Qy          65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
               ||||||||||||||||||||||||  | ||||||||||||| :|| |||| |||||||||
Db          92 REQHLYYQDQLLPVSRIIVHPQFYIIQTGADIALLELEEPVNISSRVHTVMLPPASETFP 151

Qy         125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               ||||||||||||||||| ||||||||||||||||||||||||||||||||||||:|||||
Db         152 PGMPCWVTGWGDVDNDEPLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIIRDDML 211

Qy         185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
               |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db         212 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWDEGCAQPNRPGIYTRVTYYLDWIHHY 271

Qy         245 VPKKP 249
               |||||
Db         272 VPKKP 276


RESULT 4
Q7YS62
ID   Q7YS62       PRELIMINARY;      PRT;   275 AA.
AC   Q7YS62;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tryptase (EC 3.4.21.59).
GN   MTC1.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Dacre K.J., McAleese S., Pemberton A., McGorum B.C.;
RT   "cDNA sequence of two equine mast cell proteases and the differential
RT   expression of tryptase and equine mast cell protease-1 in equine
RT   tissues.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Dacre K.J.;
RT   "The role of mast cells and mast cell proteases in equine heaves.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ515902; CAD56807.1; -.
KW   Hydrolase.
SQ   SEQUENCE   275 AA;  30476 MW;  718F57A4DBEE21ED CRC64;

  Query Match             78.4%;  Score 1092.5;  DB 6;  Length 275;
  Best Local Similarity   77.8%;  Pred. No. 7.1e-99;
  Matches  193;  Conservative   19;  Mismatches   35;  Indels    1;  Gaps    1;

Qy           1 LEKR-IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAA 59
               ||:  |||||||  |||||||||| :   || ||||||||||||||||| |||||::|
Db          26 LEREGIVGGQEASGSKWPWQVSLRKNTEYWKHFCGGSLIHPQWVLTAAHCVGPDIEDFRD 85

Qy          60 LRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPA 119
               :|||||||||||:|||||||||: || :|| : |||||||||::|| :||||  ||||||
Db          86 IRVQLREQHLYYRDQLLPVSRILPHPYYYTVENGADIALLELQDPVNISSHVQVVTLPPA 145

Qy         120 SETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIV 179
               ||||||| ||||||||||||   ||||||||:|||||:|| :|| ||| |  |||::|||
Db         146 SETFPPGTPCWVTGWGDVDNGVSLPPPFPLKEVKVPIVENSVCDRKYHTGVSTGDNIRIV 205

Qy         180 RDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLD 239
               : ||||||| | ||||||||||||||| |||||||||||   ||||||||||||||||||
Db         206 QADMLCAGNRRHDSCQGDSGGPLVCKVKGTWLQAGVVSWANSCAQPNRPGIYTRVTYYLD 265

Qy         240 WIHHYVPK 247
               ||: ||||
Db         266 WIYQYVPK 273
```

```
RESULT 5
Q8HYJ2
ID   Q8HYJ2       PRELIMINARY;      PRT;   271 AA.
AC   Q8HYJ2;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tryptase.
GN   BLT.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RA   Gambacurta A., Fiorucci L.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF515641; AAN77096.1; -.
DR   PIR; S24190; S24190.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   Pfam; PF00089; trypsin; 1.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
SQ   SEQUENCE   271 AA;  29902 MW;  1B68B5A188A80CFA CRC64;

  Query Match             75.4%;  Score 1051;  DB 6;  Length 271;
  Best Local Similarity   75.9%;  Pred. No. 8.2e-95;
  Matches  186;  Conservative   22;  Mismatches   37;  Indels    0;  Gaps    0;

Qy           5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
               |||||||| |:|||||||||   || | ||||||||||||||| ||||:|   :  ||||
Db          27 IVGGQEAPGSRWPWQVSLRVSRRYWRHHCGGSLIHPQWVLTAAHCVGPEVHGPSYFRVQL 86

Qy          65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
               |||||||||||||:|||| || :|: : |||||||||:: | :| ||  ||||| |||||
Db          87 REQHLYYQDQLLPISRIIPHPNYYSVKNGADIALLELDKLVNISWHVQLVTLPPESETFP 146

Qy         125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               ||  |||||||:|||  ||||||||||||||::|| :|| ||| |  |||:| ||::| |
Db         147 PGTQCWVTGWGNVDNGRRLPPPFPLKQVKVPVVENSVCDRKYHSGLSTGDNVSIVQEDNL 206

Qy         185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
               |||:: |||||||||||||||||||||||||||||:|||:||||||||||| |||||| |
Db         207 CAGDSGRDSCQGDSGGPLVCKVNGTWLQAGVVSWGDGCAKPNRPGIYTRVTSYLDWIHQY 266

Qy         245 VPKKP 249
               ||: |
Db         267 VPQGP 271


RESULT 6
Q921N4
ID   Q921N4       PRELIMINARY;      PRT;   273 AA.
AC   Q921N4;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mast cell protease 7.
GN   MCPT7.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RA   Strausberg R.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; BC011328; AAH11328.1; -.
DR   HSSP; P00761; 1AN1.
DR   MGD; MGI:96943; Mcpt7.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   273 AA;  30332 MW;  51417476514035BE CRC64;

  Query Match             75.1%;  Score 1046;  DB 11;  Length 273;
  Best Local Similarity   76.1%;  Pred. No. 2.6e-94;
  Matches  185;  Conservative   17;  Mismatches   41;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |||||||  :||||||||| :  |||||||||||||||||||| |||||| |   :||||
Db         29 IVGGQEAHGNKWPWQVSLRANDTYWMHFCGGSLIHPQWVLTAAHCVGPDVADPNKVRVQL 88

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |:|:||| | |: ||:|| || ||  | |||||||:|  ||  :| :|| | |||||||||
Db         89 RKQYLYYHDHLMTVSQIITHPDFYIVQDGADIALLKLTNPVNISDYVHPVPLPPASETFP 148

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
               |  |||||||::||   ||||||||:|:|||:|||:|| ||| |  |||:| |||||||
Db        149 SGTLCWVTGWGNIDNGVNLPPPFPLKEVQVPIIENHLCDLKYHKGLITGDNVHIVRDDML 208

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              ||||   |||||||||||||||  |||||||||||||||||||||||||||||||||| |
Db        209 CAGNEGHDSCQGDSGGPLVCKVEDTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHRY 268

Qy        245 VPK 247
              |||
Db        269 VPK 271

RESULT 7
Q96RZ7
ID   Q96RZ7      PRELIMINARY;      PRT;   233 AA.
AC   Q96RZ7;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mast cell tryptase beta III.
GN   TRYPTASEB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21096910; PubMed=11157797;
RA   Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA   Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA   Higgs D.R.;
RT   "Sequence, structure and pathology of the fully annotated terminal 2
RT   Mb of the short arm of human chromosome 16.";
RL   Hum. Mol. Genet. 10:339-352(2001).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AE006466; AAK61270.1; -.
DR   HSSP; P00761; 1AN1.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
```

```
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   233 AA;  25874 MW;  246A96C71B2CE6D0 CRC64;

  Query Match             73.4%;  Score 1022;  DB 4;  Length 233;
  Best Local Similarity   97.4%;  Pred. No. 4.8e-92;
  Matches  186;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              ||||||||||||||||||||   |||||||||||||||||||| ||||||||||||||||
Db         31 IVGGQEAPRSKWPWQVSLRVRDRYWMHFCGGSLIHPQWVLTAAHCVGPDVKDLAALRVQL 90

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              ||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db         91 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVNVSSHVHTVTLPPASETFP 150

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        151 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 210

Qy        185 CAGNTRRDSCQ 195
              |||||||||||
Db        211 CAGNTRRDSCQ 221

RESULT 8
Q9XSM1
ID   Q9XSM1      PRELIMINARY;      PRT;   273 AA.
AC   Q9XSM1;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tryptase (EC 3.4.21.59).
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Bone marrow;
RX   MEDLINE=20308142; PubMed=10848900;
RA   Pemberton A.D., McAleese S.M., Huntley J.F., Collie D.D.S.,
RA   Scudamore C.L., McEuen A.R., Walls A.F., Miller H.R.P.;
RT   "cDNA sequence of two sheep mast cell tryptases and the differential
RT   expression of tryptase and sheep mast cell proteinase-1 in lung,
RT   dermis and gastrointestinal tract.";
RL   Clin. Exp. Allergy 30:818-883(2000).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; Y18223; CAB41988.1; -.
DR   HSSP; P20231; 1A0L.
DR   MEROPS; S01.118; -.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0030019; F:tryptase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
```

```
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
FT   CHAIN        29    273       TRYPTASE.
SQ   SEQUENCE   273 AA;  30156 MW;  88A19DC779053898 CRC64;

  Query Match             73.2%;  Score 1019;  DB 6;  Length 273;
  Best Local Similarity   71.8%;  Pred. No. 1.1e-91;
  Matches  176;  Conservative   29;  Mismatches   40;  Indels    0;  Gaps    0;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQL 64
              |:||:||| |:|||||||||   || | ||||||||||||||| |:||:::: :  ||||
Db         29 IIGGKEAPGSRWPWQVSLRVRDQYWRHQCGGSLIHPQWVLTAAHCIGPELQEPSDFRVQL 88

Qy         65 REQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFP 124
              |||||||||:|||:||:| || :|  : |||||||:||||| :| ||  |||||||||||
Db         89 REQHLYYQDRLLPISRVIPHPHYYMVENGADIALLQLEEPVSISCHVRPVTLPPASETFP 148

Qy        125 PGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDML 184
              ||  |||||||:|||   ||||:|||||||||:|| :|| ||| |  |   | ||::| |
Db        149 PGSQCWVTGWGNVDNGRPLPPPYPLKQVKVPIVENSVCDWKYHSGLSTDYSVPIVQEDNL 208

Qy        185 CAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHY 244
              |||:  |||||||||||||||||||||||||||:||| |: ||:|||:| |||||| |
Db        209 CAGDGGRDSCQGDSGGPLVCKVNGTWLQAGVVSWGDGCANPDYPGVYTRITSYLDWIHQY 268

Qy        245 VPKKP 249
              ||::|
Db        269 VPQEP 273

RESULT 9
Q29464
ID   Q29464      PRELIMINARY;      PRT;   237 AA.
AC   Q29464;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tryptase (EC 3.4.21.59) (Fragment).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver capsule;
RX   MEDLINE=96203914; PubMed=8620861;
RA   Pallaoro M., Gambacurta A., Fiorucci L., Mignogna G., Barra D.,
RA   Ascoli F.;
RT   "cDNA cloning and primary structure of tryptase from bovine mast cells
RT   and evidence of the expression of bovine pancreatic trypsin inhibitor
RT   mRNA in the same cells.";
RL   Eur. J. Biochem. 237:100-105(1996).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; X94982; CAA64438.1; -.
DR   PIR; S68702; S68702.
DR   HSSP; P20231; 1A0L.
DR   MEROPS; S01.118; -.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0030019; F:tryptase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
KW   Hydrolase; Protease; Serine protease.
FT   NON_TER       1      1
```

```
SQ   SEQUENCE   237 AA;  26550 MW;  EA537A75294EFFBA CRC64;

  Query Match             72.1%;  Score 1005;  DB 6;  Length 237;
  Best Local Similarity   75.4%;  Pred. No. 2.3e-90;
  Matches  178;  Conservative   21;  Mismatches   37;  Indels    0;  Gaps    0;

Qy         14 SKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQLREQHLYYQD 73
              |:|||||||||   || | ||||||||||||||| ||||:|     :|||||||||||||
Db          2 SQWPWQVSLRVSRRYWRHHCGGSLIHPQWVLTAAHCVGPEVHGPSYFRVQLREQHLYYQD 61

Qy         74 QLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASETFPPGMPCWVTG 133
              ||||:|||| ||  |: : |||||||||:: | :| ||  ||||| |||||||  |||||
Db         62 QLLPISRIIPHPNCYSVKNGADIALLELDKLVNISWHVQPVTLPPESETFPPGTQCWVTG 121

Qy        134 WGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDS 193
              ||:|||  ||||||||||||||::|| :|| ||| |  |||:| |||:||||||:: |:
Db        122 WGNVDNGRRLPPPFPLKQVKVPVVENSVCDRKYHSGLSTGDNVPIVREDMLCAGDSGRNF 181

Qy        194 CQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP 249
              ||||||||||||||||||||||||||:|||:||||||||||| |||||| |||: |
Db        182 CQGDSGGPLVCKVNGTWLQAGVVSWGDGCAKPNRPGIYTRVTSYLDWIHQYVPQGP 237

RESULT 10
Q80UR4
ID   Q80UR4      PRELIMINARY;      PRT;   318 AA.
AC   Q80UR4;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Mast cell protease-11.
GN   MCP11.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Bone-marrow;
RA   Wong G.W., Yasuda S., Li L., Stevens R.L.;
RT   "Cloning and characterization of mouse mast cell protease-11 (mMCP-
RT   11).";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY261775; AAP21675.1; -.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
KW   Protease.
SQ   SEQUENCE   318 AA;  35037 MW;  6F69B6653FF54FEA CRC64;

  Query Match             48.7%;  Score 678;  DB 11;  Length 318;
  Best Local Similarity   52.4%;  Pred. No. 3.9e-58;
  Matches  130;  Conservative   37;  Mismatches   75;  Indels    6;  Gaps    2;

Qy          5 IVGGQEAPRSKWPWQVSLRVHG---PYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALR 61
              ||||     |::|||||||::      | | ||||||||||||||| || |   :   :|
Db         35 IVGGCPVSASRFPWQVSLRLYDMEHSRWEHECGGSLIHPQWVLTAAHCVRPKEVEAYGVR 94

Qy         62 VQLREQHLYYQDQLLPVSRIIVHPQF---YTAQIGADIALLELEEPVKVSSHVHTVTLPP 118
              ||: :  ||  |||: | :|| ||:|    :|: |||||||:|:  | :| ||: |:||
Db         95 VQVGQLRLYENDQLMKVVKIIRHPKFSEKLSARGGADIALLKLDTRVVLSEHVYPVSLPA 154

Qy        119 ASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRI 178
              ||        ||| ||| ::|   ||||: |::| |||:||: |: ||   : :    ||
```

```
Db        155 ASLRISSKKTCWVAGWGVIENYMPLPPPYHLREVAVPIVENNDCEQKYQTNSSSDSTTRI 214

Qy        179 VRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYL 238
              ::|||||||   ||||: ||||||||: | :|:| |||||| ||  |: ||:||||  |:
Db        215 IKDDMLCAGKEGRDSCKADSGGPLVCRWNCSWVQVGVVSWGIGCGLPDFPGVYTRVMSYV 274

Qy        239 DWIHHYVP 246
               ||  |||
Db        275 SWIKCYVP 282


RESULT 11
Q8SQ44
ID   Q8SQ44      PRELIMINARY;      PRT;   277 AA.
AC   Q8SQ44;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tryptase precursor.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RA   Sato M., Yoshida S., Iida K., Tomozawa T., Kido H., Yamashita M.;
RT   "A Novel Influenza A Virus Activating Enzyme from Porcine Lungs:
RT   Purification and Characterization.";
RL   Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB080655; BAB85761.1; -.
DR   HSSP; P00761; 1AN1.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease; Signal.
FT   SIGNAL        1     30       POTENTIAL.
FT   CHAIN        31    277       TRYPTASE.
SQ   SEQUENCE   277 AA;  30407 MW;  35FC8B569C0ED4C5 CRC64;

  Query Match             46.7%;  Score 651;  DB 6;  Length 277;
  Best Local Similarity   50.6%;  Pred. No. 1.5e-55;
  Matches  126;  Conservative   39;  Mismatches   76;  Indels    8;  Gaps    3;

Qy          5 IVGGQEAPRSKWPWQVSL---RVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALR 61
              |||| :    |:||||||   ||    ||| |||||:||||||||| ||         :|
Db         31 IVGGHDVSTWKYPWQVSLRSYRVKLGRWMHHCGGSLVHPQWVLTAAHCVDSHNLKPQDVR 90

Qy         62 VQLREQHLYYQDQLLPVSRIIVHPQFY-TAQIGADIALLELEEPVKVSSHVHTVTLPPAS 120
              ||: :  ||  |||  | :|| ||::   |: |||||||:|| |: :|: |: |  | |:
Db         91 VQVGQLKLYDGDQLTKVKQIIRHPKYLGFAKGGADIALLQLEAPLTLSARVNVVGTPSAT 150

Qy        121 ETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVR 180
                 | |  | |||||:: ::  ||||: |::|:|||: | :|:  |     || : : ::
Db        151 LKVPKGKRCSVTGWGNIKHNLSLPPPYHLQEVEVPIVANKVCNKHYR----TGPNSKPIK 206

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
               ||||||:   |||||||||||:|  ||||:| |:|||| ||   | ||:| ||  |: |
Db        207 ADMLCAGSKGLDSCQGDSGGPLMCSWNGTWVQVGIVSWGRGCGLHNFPGVYIRVMSYVSW 266

Qy        241 IHHYVPKKP 249
              |: |||: |
Db        267 IYQYVPRSP 275
```

```
RESULT 12
Q99MS4
ID   Q99MS4      PRELIMINARY;      PRT;   279 AA.
AC   Q99MS4;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Implantation serine proteinase 2.
GN   ISP2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21361749; PubMed=11467974;
RA   O'Sullivan C.M., Liu S.Y., Rancourt S.L., Rancourt D.E.;
RT   "Regulation of the strypsin-related proteinase ISP2 by progesterone in
RT   endometrial gland epithelium during implantation in mice.";
RL   Reproduction 122:235-244(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Uterus;
RA   Shen Q.-X., Wang J., Huang Z.-P.;
RT   "Identification of endometrial factors involved in mouse embryo
RT   implantation.";
RL   Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AF305425; AAK15264.2; -.
DR   EMBL; AF442819; AAL38005.1; -.
DR   HSSP; P00763; 1DPO.
DR   MEROPS; S01.315; -.
DR   MGD; MGI:2149952; Isp2.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   279 AA;  30986 MW;  81C43A59935244D7 CRC64;

  Query Match             45.9%;  Score 640;  DB 11;  Length 279;
  Best Local Similarity   49.8%;  Pred. No. 1.8e-54;
  Matches  120;  Conservative   36;  Mismatches   81;  Indels    4;  Gaps    2;

Qy          5 IVGGQEAPRSKWPWQVSLRVHGPY---WMHFCGGSLIHPQWVLTAAACVGPDVKDLAALR 61
              ||||  ||: |||||||||::  |   |:| ||||:|||||||||| |:     | :  |
Db         31 IVGGHSAPQGKWPWQVSLRIYRYYWAFWVHNCGGSIIHPQWVLTAAHCIRERDADPSVFR 90

Qy         62 VQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTLPPASE 121
              ::: | :||   :|| |||:|:|| |  | :|:|:|||:|   |:   :|  | ||  |
Db         91 IRVGEAYLYGGKELLSVSRVIIHPDFVHAGLGSDVALLQLAVSVQSFPNVKPVKLPSESL 150

Qy        122 TFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGA-YTGDDVRIVR 180
                     ||||||| |     ||||: |:||:| |::| :|:  ||    :     :::
Db        151 EVTKKDVCWVTGWGAVSTHRSLPPPYRLQQVQVKIIDNSLCEEMYHNATRHRNRGQKLIL 210

Qy        181 DDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDW 240
               |||||||  :||| ||||||||| | |:|   |||||| |||  : ||:| ||  :| |
Db        211 KDMLCAGNQGQDSCYGDSGGPLVCNVTGSWTLVGVVSWGYGCALRDFPGVYARVQSFLPW 270

Qy        241 I 241
              |
```

```
Db        271 I 271


RESULT 13
Q8BJR6
ID   Q8BJR6      PRELIMINARY;      PRT;   328 AA.
AC   Q8BJR6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Similar to MARAPSIN precursor (Channel-activating protease 2-like
DE   protein) (Pancreasin).
GN   MPN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Thymus;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129S6/SvEvTac;
RA   Brathwaite M., Waeltz P., Schlessinger D., Nagaraja R.;
RT   "Genomic Sequence Analysis in the Mouse T-complex Region.";
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Bladder;
RX   MEDLINE=22439795; PubMed=12441343;
RA   Bhagwandin V.J., Hau L.W., Mallen-St  Clair J., Wolters P.J.,
RA   Caughey G.H.;
RT   "Structure and activity of human pancreasin, a novel tryptic serine
RT   peptidase expressed primarily by the pancreas.";
RL   J. Biol. Chem. 5:3363-3371(2003).
DR   EMBL; AK080281; BAC37864.1; -.
DR   EMBL; AY162410; AAO17162.1; -.
DR   EMBL; AF542056; AAO27572.1; -.
DR   MGD; MGI:2450123; Mpn.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Protease.
SQ   SEQUENCE   328 AA;  35789 MW;  DC0B20F1AB3EB840 CRC64;

  Query Match             42.8%;  Score 596.5;  DB 11;  Length 328;
  Best Local Similarity   45.1%;  Pred. No. 4e-50;
  Matches  114;  Conservative   48;  Mismatches   74;  Indels   17;  Gaps    5;

Qy          4 RIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKD------- 56
              |:|||: |   :||||||:: :|   :|||||||| | |||||| |   :  |
Db         37 RMVGGENALEGEWPWQVSIQRNG---IHFCGGSLIAPTWVLTAAHCFS-NTSDISIYQVL 92

Qy         57 LAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTVTL 116
              | ||::|    |  |    :|| ::  :||:      ||:||:||: ||  ::::  | |
Db         93 LGALKLQQPGPHALY----VPVKQVKSNPQYQGMASSADVALVELQGPVTFTNYILPVCL 148
```

```
Qy        117 PPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDV 176
              |  |  |  || |||||||     :||| |  |::: |||::   |:  |:    :   :
Db        149 PDPSVIFESGMNCWVTGWGSPSEQDRLPNPRVLQKLAVPIIDTPKCNLLYNKDVESDFQL 208

Qy        177 RIVRDDMLCAG--NTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRV 234
              : ::|||||||    ::|:|:||||||||| |: :|:||||:|||||||: ||||:| ||
Db        209 KTIKDDMLCAGFAEGKKDACKGDSGGPLVCLVDQSWVQAGVISWGEGCARRNRPGVYIRV 268

Qy        235 TYYLDWIHHYVPK 247
              | :  |||  :|:
Db        269 TSHHKWIHQIIPE 281


RESULT 14
Q96RZ8
ID   Q96RZ8      PRELIMINARY;      PRT;   321 AA.
AC   Q96RZ8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   HS transmembrane tryptase, gene name TMT, AF175522_1.
GN   TRPA.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21096910; PubMed=11157797;
RA   Daniels R.J., Peden J.F., Lloyd C., Horsley S.W., Clark K.,
RA   Tufarelli C., Kearney L., Buckle V.J., Doggett N.A., Flint J.,
RA   Higgs D.R.;
RT   "Sequence, structure and pathology of the fully annotated terminal 2
RT   Mb of the short arm of human chromosome 16.";
RL   Hum. Mol. Genet. 10:339-352(2001).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
DR   EMBL; AE006466; AAK61269.1; -.
DR   HSSP; P00761; 1AN1.
DR   GO; GO:0016021; C:integral to membrane; IEA.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
KW   Hydrolase; Protease; Serine protease; Transmembrane.
SQ   SEQUENCE   321 AA;  33829 MW;  FFF5089EDC4FC73D CRC64;

  Query Match             42.3%;  Score 589.5;  DB 4;  Length 321;
  Best Local Similarity   48.4%;  Pred. No. 1.9e-49;
  Matches  118;  Conservative   33;  Mismatches   80;  Indels   13;  Gaps    6;

Qy          4 RIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAACVGPDVKDLAALRVQ 63
              |||||  ||   |||| |||:     :| |||||: |||||||| |    :   :  :|
Db         37 RIVGGHAAPAGAWPWQASLRLRR---VHVCGGSLLSPQWVLTAAHCFSGSLNS-SDYQVH 92

Qy         64 LREQHLYYQDQLLPVSRIIVHPQFYTAQIG--ADIALLELEEPVKVSSHVHTVTLPPASE 121
              | |  :        | :||:|    : | |   ||||:||  || :|| :  | || ||:
Db         93 LGELEITLSPHFSTVRQIILHSS-PSGQPGTSGDIALVELSVPVTLSSRILPVCLPEASD 151

Qy        122 TFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGDDVRIVRD 181
               | ||: |||||||     | ||||: |::||| :::   |        | |    |::
Db        152 DFCPGIRCWVTGWGYTREGEPLPPPYSLREVKVSVVDTETCRRD-----YPGPGGSILQP 206

Qy        182 DMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYTRVTYYLDWI 241
              |||||     |:|| ||||||||:||| |:|||:||||||| :|||||:||||  |::||
Db        207 DMLCARGP-GDACQDDSGGPLVCQVNGAWVQAGIVSWGEGCGRPNRPGVYTRVPAYVNWI 265
```

```
Qy        242 HHYV 245
                ::
Db        266 RRHI 269


RESULT 15
Q80WM7
ID   Q80WM7       PRELIMINARY;      PRT;   277 AA.
AC   Q80WM7;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Tryptase-6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Testis;
RA   Wong G.W., Yasuda S., Li L., Stevens R.L.;
RT   "Cloning and characterization of mouse tryptase-6 (mT6).";
RL   Submitted (MAR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY262280; AAP20885.1; -.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR009003; Cys_Ser_trypsin.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
SQ   SEQUENCE   277 AA;  29887 MW;  3310C067573E5E74 CRC64;

  Query Match             42.2%;  Score 587.5;  DB 11;  Length 277;
  Best Local Similarity   46.2%;  Pred. No. 2.5e-49;
  Matches  115;  Conservative   34;  Mismatches   85;  Indels   15;  Gaps    5;

Qy          1 LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAAC----VGPDVKD 56
              :  |||||::|   :|||| |::  |    | |||||| |||||||  |    | |
Db         30 MSSRIVGGRDAQDGEWPWQTSIQHRG---AHVCGGSLIAPQWVLTAGHCFPRRVWPSEYS 86

Qy         57 --LAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVKVSSHVHTV 114
                | || : :|  |    : |:|| |::: | :   :   |:|||:|  || :|: :  |
Db         87 VLLGALSLDVRSSH----ELLVPVLRVLLPPDYSEDEARGDLALLQLRHPVSLSTRIQPV 142

Qy        115 TLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDAKYHLGAYTGD 174
               ||      ||| |||||||| :     ||   ||: |:||::::  ||  ||:||
Db        143 CLPAPGSHPPPGSPCWVTGWGSLSPGVPLPKGRPLQGVRVPLLDSRACDRLYHVGANVPQ 202

Qy        175 DVRIVRDDMLCAGNTR--RDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGCAQPNRPGIYT 232
                |||    ||||  |  :|:||||||||| |  :| |:   |||||:||| |||||:||
Db        203 GERIVLPGNLCAGYRRGHKDACQGDSGGPLTCMESGHWVLVGVVSWGKGCALPNRPGVYT 262

Qy        233 RVTYYLDWI 241
               |  |  ||
Db        263 NVAKYSPWI 271

Search completed: July 22, 2004, 15:25:07
Job time : 85 secs
```

```
                      GenCore version 5.1.6
              Copyright (c) 1993 - 2004  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:          July 22, 2004, 15:33:07 ; Search time 3069 Seconds
                                           (without alignments)
                                           2422.836 Million cell updates/sec

Title:           US-09-598-982-21
Perfect score:   1393
Sequence:        1 LEKRIVGGQEAPRSKWPWQV..........IYTRVTYYLDWIHHYVPKKP 249

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters:      55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09598982/runat_20072004_064410_6500/app_query.fasta_1.391
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09598982_@CGN_1_1_6425_@runat_20072004_064410_6500 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  em_gss_hum:*
                 18:  em_gss_inv:*
                 19:  em_gss_pln:*
                 20:  em_gss_vrt:*
                 21:  em_gss_fun:*
                 22:  em_gss_mam:*
                 23:  em_gss_mus:*
                 24:  em_gss_pro:*
                 25:  em_gss_rod:*
                 26:  em_gss_phg:*
                 27:  em_gss_vrl:*
                 28:  gb_gss1:*
```

```
                 29:  gb_gss2:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                      SUMMARIES
                        %
Result                Query
  No.     Score   Match Length DB  ID                          Description
-----------------------------------------------------------------------------------
     1    1263    90.7   1030  12  BM919268                    BM919268 AGENCOURT
     2  1109.5    79.6    904  13  BQ720404                    BQ720404 AGENCOURT
     3    1062    76.2    589  13  BU071733                    BU071733 im33a09.y
     4    1060    76.1    619  14  CD671851                    CD671851 fg07g11.y
     5    1024    73.5   1000  12  BI413250                    BI413250 602986390
     6    1023    73.4    930  14  CB203717                    CB203717 AGENCOURT
     7     972    69.8    635  14  CD671891                    CD671891 fg08b10.y
     8     966    69.3    642  14  CD672689                    CD672689 fg15a06.y
     9     960    68.9    722  13  BX091892                    BX091892 BX091892
    10   946.5    67.9    811  12  BG697540                    BG697540 602660760
    11     915    65.7    502  13  BQ082929                    BQ082929 K-EST0144
    12     905    65.0    581  13  BU787772                    BU787772 io34c08.y
    13     891    64.0    736  12  BG542089                    BG542089 602571309
c   14     855    61.4    929  14  CF583026                    CF583026 AGENCOURT
    15   845.5    60.7    806  14  CF553058                    CF553058 AGENCOURT
    16     832    59.7    917  13  BQ721078                    BQ721078 AGENCOURT
c   17     800    57.4    686  13  BU073263                    BU073263 im33a09.x
    18     800    57.4    859  14  CF583025                    CF583025 AGENCOURT
    19     789    56.6    639  9  AA131142                    AA131142 zl31b01.r
    20     785    56.4    678  12  BG483591                    BG483591 602503309
    21     773    55.5    671  14  CB840425                    CB840425 M15E-0715
    22     715    51.3    474  10  AW823937                    AW823937 uf61e03.y
    23     690    49.5    558  12  BI790801                    BI790801 id09c06.y
    24     681    48.9    512  10  BF724180                    BF724180 bx01h02.y
c   25     680    48.8    699  14  CA438721                    CA438721 UI-H-DT1-
    26     678    48.7   1219  11  AK081986                    AK081986 Mus muscu
c   27     672    48.2    690  14  CA439686                    CA439686 UI-H-DT1-
c   28     662    47.5    688  12  BQ003361                    BQ003361 UI-H-EI1-
c   29     657    47.2    693  12  BM991728                    BM991728 UI-H-DF1-
c   30     654    46.9    692  12  BM989945                    BM989945 UI-H-DI0-
    31     653    46.9   1928  11  AK040500                    AK040500 Mus muscu
c   32     644    46.2    371  10  BF359130                    BF359130 QV3-ET006
c   33     638    45.8    682  14  CA425354                    CA425354 UI-H-DF0-
    34   629.5    45.2    676  12  BG964077                    BG964077 602828831
c   35     624    44.8    376  10  BF850308                    BF850308 CM3-EN007
    36     623    44.7    660  14  CB589106                    CB589106 AGENCOURT
    37     622    44.7    620  9  AA049080                    AA049080 mj50f09.r
    38     618    44.4    411  10  BE862270                    BE862270 UI-M-BH0-
    39     599    43.0    525  12  BI775716                    BI775716 468400 MA
    40     598    42.9    545  12  BI681216                    BI681216 460544 MA
    41   596.5    42.8   1150  11  AK080281                    AK080281 Mus muscu
    42   592.5    42.5    730  29  CC546157                    CC546157 CH240_429
    43     584    41.9    581  14  CA867461                    CA867461 ir28f02.y
    44     566    40.6    538  10  BE751979                    BE751979 204246 MA
c   45     560    40.2    569  12  BI287271                    BI287271 UI-R-CW0s


                                      ALIGNMENTS

RESULT 1
BM919268
LOCUS       BM919268                1030 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6715837 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748499
            5', mRNA sequence.
ACCESSION   BM919268
VERSION     BM919268.1  GI:19369647
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1030)
```

```
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE    National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL  Unpublished (1999)
COMMENT    Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: LLAM12776  row: f  column: 20
           High quality sequence stop: 719.
FEATURES             Location/Qualifiers
     source          1. .1030
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5748499"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_120"
                     /note="Organ: pooled pancreas and spleen; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of spleen and pancreas from 28 yo
                     male.  Library is oligo-dT primed and directionally cloned
                     (EcoRV  site is destroyed upon cloning).  Average insert
                     size 1.5 kb, insert size range 1-2.5 kb.  Library is
                     normalized and enriched for  full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 025. Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              4.9e-117     Length:         1030
Score:                  1263.00      Matches:        228
Percent Similarity:     93.93%       Conservative:   4
Best Local Similarity:  92.31%       Mismatches:     13
Query Match:            90.67%       Indels:         2
DB:                     12           Gaps:           0

US-09-598-982-21 (1-249) x BM919268 (1-1030)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        123 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 182

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        183 CGCGACCGATACTGGATGCACTTCTGTGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 242

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   |||:::||||||||||||||||||||||||   |||||||||||||||
Db        243 ACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTG 302

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        303 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 362

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db        363 CCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCC 422

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::||||||   ||||||||||||   |||||||||||||||||||||||||||
Db        423 GTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCC 482

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        483 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCA 542

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        543 CCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 602

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db        603 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTG 662

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        663 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCTGGAGGGCCCCTGGTGTGC 722

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrpGly--GluGlyCysAlaG 224
              |||||||||||||||||||||||||||||||||||||||||||||     ||||||||||
Db        723 AAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGGACGAGGGGCTGTGCCC 782

Qy        224 lnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisT 244
              |||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        783 AGCCCCACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 842

Qy        244 yrValProLysLysPro 249
              |||||||||||||||||
Db        843 ATGTCCCCAAAAAGCCG 859

RESULT 2
BQ720404
LOCUS       BQ720404                 904 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8477973 Lupski_sympathetic_trunk Homo sapiens cDNA clone
            IMAGE:6196776 5', mRNA sequence.
ACCESSION   BQ720404
VERSION     BQ720404.1  GI:21859301
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13605  row: m  column: 01
            High quality sequence stop: 587.
FEATURES             Location/Qualifiers
     source          1. .904
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6196776"
                     /sex="male"
                     /tissue_type="sympathetic trunk"
                     /dev_stage="adult, 16 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_sympathetic_trunk"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.9 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine); available through Life
                     Technologies."
ORIGIN

Alignment Scores:
```

```
Pred. No.:              1.39e-101     Length:         904
Score:                  1109.50       Matches:        207
Percent Similarity:     88.24%        Conservative:   3
Best Local Similarity:  86.97%        Mismatches:     17
Query Match:            79.65%        Indels:         11
DB:                     13            Gaps:           1

US-09-598-982-21 (1-249) x BQ720404 (1-904)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 159

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 219

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||                           |||||||||
Db        220 ACCGCAGCGCACTGCGTGGGACCG---------------------------GTGCAACTG 252

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        253 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 312

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        313 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 372

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        373 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 432

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        433 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 492

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        493 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 552

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        553 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 612

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        613 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGGGC 672

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyVal-ValSerTrpGlyGluGlyCysAlaGl 224
                    ||||||   ||||||   ||||||||| |||:::||||||||||||   |||::
Db        673 AACGCGAATGGCCCCTGGCTGGCAGCGGGCGTGGGTCACCTGGGGCGAGGGCTGGGCCCG 732

Qy        224 nProAsnArgProGlyIleTyrThr-ArgValThrTyrTyrLeuAspTrp 240
              :|||||||||   |||   |||||| |||         |||   :::|||
Db        733 GCCCAACCGGGCGGGCTTCTACACCCCGGGCCCCGTCCTACTGGGAATGG 782

RESULT 3
BU071733
LOCUS       BU071733                 589 bp    mRNA    linear   EST 27-AUG-2002
DEFINITION  im33a09.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6036761
            5' similar to SW:TRYB_HUMAN P20231 BETA-TRYPTASE PRECURSOR ;, mRNA
            sequence.
ACCESSION   BU071733
VERSION     BU071733.1  GI:22512922
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 589)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: im33a09.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 433.
FEATURES             Location/Qualifiers
     source          1..589
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6036761"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:              4.62e-97      Length:         589
Score:                  1062.00       Matches:        192
Percent Similarity:     99.48%        Conservative:   0
Best Local Similarity:  99.48%        Mismatches:     1
Query Match:            76.24%        Indels:         0
DB:                     13            Gaps:           0

US-09-598-982-21 (1-249) x BU071733 (1-589)

Qy         53 AspValLysAspLeuAlaAlaLeuArgValGlnLeuArgGluGlnHisLeuTyrTyrGln 72
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         10 GACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAG 69

Qy         73 AspGlnLeuLeuProValSerArgIleIleValHisProGlnPheTyrThrAlaGlnIle 92
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         70 GACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATC 129

Qy         93 GlyAlaAspIleAlaLeuLeuGluLeuGluGluProValLysValSerSerHisValHis 112
              |||||||||||||||||||||||||||||||||||||||   ||||||||||||||||||
Db        130 GGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCAC 189

Qy        113 ThrValThrLeuProProAlaSerGluThrPheProProGlyMetProCysTrpValThr 132
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        190 ACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACT 249

Qy        133 GlyTrpGlyAspValAspAsnAspGluArgLeuProProProPheProLeuLysGlnVal 152
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        250 GGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTG 309
```

```
Qy         153 LysValProIleMetGluAsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThr 172
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         310 AAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACG 369

Qy         173 GlyAspAspValArgIleValArgAspAspMetLeuCysAlaGlyAsnThrArgArgAsp 192
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         370 GGAGACGACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGAC 429

Qy         193 SerCysGlnGlyAspSerGlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGln 212
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         430 TCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAG 489

Qy         213 AlaGlyValValSerTrpGlyGluGlyCysAlaGlnProAsnArgProGlyIleTyrThr 232
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         490 GCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACC 549

Qy         233 ArgValThrTyrTyrLeuAspTrpIleHisHisTyrVal 245
               |||||||||||||||||||||||||||||||||||||||
Db         550 CGTGTCACCTACTACTTGGACTGGATCCACCACTATGTC 588

RESULT 4
CD671851
LOCUS       CD671851                 619 bp    mRNA    linear   EST 24-JUN-2003
DEFINITION  fg07g11.y1 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
            fg07g11 5', mRNA sequence.
ACCESSION   CD671851
VERSION     CD671851.1  GI:32173582
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 619)
  AUTHORS   Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of adult human iris for the NEIBank
            Project: steroid-response factors and similarities with retinal
            pigment epithelium
  JOURNAL   Mol. Vis. 8 (4), 185-195 (2002)
  MEDLINE   22103462
   PUBMED   12107412
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 07  row: g  column: 11
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..619
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fg07g11"
                     /tissue_type="Iris"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Iris cDNA (Normalized): fg"
                     /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
                     library (bx) was normalized by self-subtraction. One
                     portion of double stranded plasmid DNA representing the
                     library was linearized by NotI. This NotI digested library
                     was used as a template for biotinylated RNA synthesis
                     using SP6 RNA polymerase. Another portion of the double
                     stranded plasmid library was converted to single-stranded
                     circles in vitro using Gene II and Exonuclease III (Life
                     Technologies). Single-stranded DNA (1 mg) was hybridized
                     (C0t 500) with 41 mg of Bio-RNA and vector blocking
                     oligonucleotides. The hybridized Bio-RNA/ss-circles were
                     removed by streptavidin:phenol extraction. EST analysis
                     was performed on the library at the NIH Intramural
                     Sequencing Center(NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              7.95e-97        Length:         619
Score:                  1060.00         Matches:        190
Percent Similarity:     94.15%          Conservative:   3
Best Local Similarity:  92.68%          Mismatches:     12
Query Match:            76.09%          Indels:         0
DB:                     14              Gaps:           0

US-09-598-982-21 (1-249) x CD671851 (1-619)

Qy           8 GlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgValHisGlyPro 27
               |||||||||||||||||||||||||||||||||||||||||||||||||||
Db           3 GGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACCGA 62

Qy          28 TyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeuThrAlaAla 47
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          63 TACTGGATGCACTTCTGTGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGGCG 122

Qy          48 AlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeuArgGluGln 67
                  |||:::||||||||||||||||||||||||   ||||||||||||||||||||||||
Db         123 CACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGCAG 182

Qy          68 HisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHisProGlnPhe 87
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         183 CACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTC 242

Qy          88 TyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluProValLysVal 107
               |||      |||   |||||||||||||||||||||||||||||||||||||||   :::
Db         243 TACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATC 302

Qy         108 SerSerHisValHisThrValThrLeuProProAlaSerGluThrPheProProGlyMet 127
               ||||||   ||||||||||||   ||||||||||||||||||||||||||||||||||||
Db         303 TCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATG 362

Qy         128 ProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuProProProPhe 147
               ||||||||||||||||||||||||||||||||||||||||||   |||||||||||||||
Db         363 CCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTT 422

Qy         148 ProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAlaLysTyrHis 167
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         423 CCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCAC 482

Qy         168 LeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeuCysAlaGly 187
               |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db         483 CTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGG 542

Qy         188 AsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCysLysValAsn 207
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         543 AACACCCGGAGGGACTCATGCCAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAAT 602

Qy         208 GlyThrTrpLeuGln 212
               |||||||||||||||
Db         603 GGCACCTGGCTACAG 617

RESULT 5
BI413250
LOCUS       BI413250                1000 bp    mRNA    linear   EST 14-AUG-2001
DEFINITION  602986390F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5142478 5',
            mRNA sequence.
ACCESSION   BI413250
VERSION     BI413250.1  GI:15174173
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1000)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
```

```
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by:Incyte Genomics, Inc.
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11351  row: c  column: 23
            High quality sequence start: 32
            High quality sequence stop: 849.
FEATURES             Location/Qualifiers
     source          1. .1000
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="Czech II"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5142478"
                     /tissue_type="pooled lung tumors"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NCI_CGAP_Lu33"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
                     strand cDNA was prepared from mRNA obtained from pooled
                     lung tumors with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCTGTTTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia),  digested with Not I and cloned into the Not
                     I and Eco RI sites of  the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

Alignment Scores:
Pred. No.:              7.38e-93       Length:         1000
Score:                  1024.00        Matches:        188
Percent Similarity:     84.15%         Conservative:   19
Best Local Similarity:  76.42%         Mismatches:     35
Query Match:            73.51%         Indels:         4
DB:                     12             Gaps:           0

US-09-598-982-21 (1-249) x BI413250 (1-1000)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||   ||||||      ||||||||||||||||||||||||||||||
Db         75 ATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGATTT 134

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       ||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        135 AAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGCTC 194

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||   :::|||               ||||||||||||
Db        195 ACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGCTT 254

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              |||||||||:::|||||||||   ||||||||||||   ::::::||||||:::||||||
Db        255 CGTGAGCAGTATCTATACTATGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGCAC 314

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              |||   :::|||||||||:::   |||||||||:::||||||||||||||||||   |||
Db        315 CCCCACTATTACACGGCCGAGGGTGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCCCT 374

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||:::|||:::|||   ::::::|||||||||||||||||||||||||||
Db        375 GTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCCCC 434

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||      ||||||||||||||||||||||||:::||||||||||||   ||||||
Db        435 CCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCCCA 494
```

```
Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||:::|||||||||||||||||||||||||||:::||||||   :::||||||
Db        495 CCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACCGG 554

Qy        165 Lys-TyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLe 184
              |||    |||   |||   |||||||||||||||      ||||||   |||   |||||
Db        555 AAGTAACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGCT 614

Qy        184 uCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCy 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        615 GTGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCTG 674

Qy        204 sLysValAsnGlyThrTrpLeuGlnAlaGlyValValSer-TrpGlyGluGlyCysAlaG 224
              |||||||   |||||||||||||||||||||||||||::: |||||||||||||||||||
Db        675 CAAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCACGCTGGGGTGAGGGCTGCGCAC 734

Qy        224 ln-ProAsnArgProGlyIleTyrThrArgValThr-TyrTyrLeuAspTrpIleHisHi 243
              || ||||||:::|||||||||||||||||||||||| |||||||||||||||||||||
Db        735 AGCCCAAACAAGCCTGGCATCTACACCCGGGTGACCATACTACTTAGACTGGATCCACGC 794

Qy        243 sTyrValPro 246
               |||||||||
Db        795 TTATGTCCCT 804

RESULT 6
CB203717
LOCUS       CB203717                 930 bp    mRNA    linear   EST 05-FEB-2003
DEFINITION  AGENCOURT_11288628 NIH_MGC_135 Mus musculus cDNA clone
            IMAGE:30142408 5', mRNA sequence.
ACCESSION   CB203717
VERSION     CB203717.1  GI:28240351
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. David Rowe
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM0051  row: j  column: 17
            High quality sequence stop: 681.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30142408"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_135"
                     /note="Vector: pCMVSport6.1; Site_1: EcoRV; Site_2: NotI;
                     Normalized full-length enriched library from pooled mouse
                     embryonic limb, maxilla and mandible, day 12.5, 13.5,
                     14.5, and 15.5 (size selected for the 0.5-1 kb fragments)
                     Cloned directionally, priming method: Oligo-dT. cDNA
                     enrichment: >1k bp, Average insert size 1.6k bp.
                     Normalization (Cot value): 7.5 kb. Priming sequence:
                     5'GACTAGTTCTAGATCGCGAGCGGCCGCCC(T)3' Tissue contributed
                     by, David Rowe. Library constructed by ResGen, Invitrogen
                     Corp."
ORIGIN

Alignment Scores:
Pred. No.:              8.33e-93       Length:         930
```

```
Score:                    1023.00       Matches:        188
Percent Similarity:       83.81%        Conservative:   19
Best Local Similarity:    76.11%        Mismatches:     36
Query Match:              73.44%        Indels:         4
DB:                       14            Gaps:           0

US-09-598-982-21 (1-249) x CB203717 (1-930)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||   ||||||      ||||||||||||||||||||||||||||||   
Db         42 ATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGATTT 101

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       ||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        102 AAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGCTC 161

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||   :::|||               ||||||||||||
Db        162 ACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGCTT 221

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              |||||||||:::|||||||||   ||||||||||||   ::::::||||||:::||||||
Db        222 CGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGCAC 281

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              |||   :::|||||||||:::   |||||||||:::||||||||||||||||||   |||
Db        282 CCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCCCT 341

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||:::|||:::|||   ::::::|||||||||||||||||||||||||||
Db        342 GTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCCCC 401

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||      ||||||||||||||||||||||||:::||||||||||||   ||||||
Db        402 CCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCCCA 461

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||:::|||||||||||||||||||||||||||:::||||||   :::||||||   
Db        462 CCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACCGG 521

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              |||||||||   |||   |||||||||||||||      ||||||   |||   ||||||
Db        522 AAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGCTG 581

Qy        185 CysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 204
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        582 TGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCTGC 641

Qy        205 LysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrp-GlyGluGlyCysAlaGl 224
              ||||||   |||||||||||||||||||||||||||||||||    ||||||||||||||
Db        642 AAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGNGTGAGGGCTGCGCACA 701

Qy        224 nProAsnArgProGlyIleTyrThrArg-ValThr-TyrTyrLeuAspTrpIleHis-Hi 243
              |||||||:::|||||||||||||||||| ||||||    ||||||||||||||||||   
Db        702 GCCCAACAAGCCTGGCATCTACACCCGGGGTGACATANCTACTTAGACTGGATCCACCNG 761

Qy        243 sTyrValProLys 247
               |||||||||:::
Db        762 CTATGTCCCTGAA 774

RESULT 7
CD671891
LOCUS       CD671891                 635 bp    mRNA    linear   EST 24-JUN-2003
DEFINITION  fg08b10.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
            fg08b10 5', mRNA sequence.
ACCESSION   CD671891
VERSION     CD671891.1  GI:32173622
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 635)
  AUTHORS   Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of adult human iris for the NEIBank
            Project: steroid-response factors and similarities with retinal
            pigment epithelium
  JOURNAL   Mol. Vis. 8 (4), 185-195 (2002)
  MEDLINE   22103462
   PUBMED   12107412
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 08  row: b  column: 10
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..635
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fg08b10"
                     /tissue_type="Iris"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Iris cDNA (Normalized): fg"
                     /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
                     library (bx) was normalized by self-subtraction. One
                     portion of double stranded plasmid DNA representing the
                     library was linearized by NotI. This NotI digested library
                     was used as a template for biotinylated RNA synthesis
                     using SP6 RNA polymerase. Another portion of the double
                     stranded plasmid library was converted to single-stranded
                     circles in vitro using Gene II and Exonuclease III (Life
                     Technologies). Single-stranded DNA (1 mg) was hybridized
                     (C0t 500) with 41 mg of Bio-RNA and vector blocking
                     oligonucleotides. The hybridized Bio-RNA/ss-circles were
                     removed by streptavidin:phenol extraction. EST analysis
                     was performed on the library at the NIH Intramural
                     Sequencing Center(NISC)."
ORIGIN

Alignment Scores:
Pred. No.:                6.71e-88      Length:         635
Score:                    972.00        Matches:        176
Percent Similarity:       98.32%        Conservative:   0
Best Local Similarity:    98.32%        Mismatches:     3
Query Match:              69.78%        Indels:         0
DB:                       14            Gaps:           0

US-09-598-982-21 (1-249) x CD671891 (1-635)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         97 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 156

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        157 CATGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 216

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
                 ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        217 CTCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 276

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        277 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 336

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        337 CCCCAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 396
```

```
Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        397 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 456

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 516

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        517 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 576

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMet 183
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        577 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATG 633

RESULT 8
CD672689
LOCUS       CD672689                 642 bp    mRNA    linear   EST 24-JUN-2003
DEFINITION  fg15a06.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
            fg15a06 5', mRNA sequence.
ACCESSION   CD672689
VERSION     CD672689.1  GI:32174420
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of adult human iris for the NEIBank
            Project: steroid-response factors and similarities with retinal
            pigment epithelium
  JOURNAL   Mol. Vis. 8 (4), 185-195 (2002)
  MEDLINE   22103462
   PUBMED   12107412
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 15  row: a  column: 06
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fg15a06"
                     /tissue_type="Iris"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Iris cDNA (Normalized): fg"
                     /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
                     library (bx) was normalized by self-subtraction. One
                     portion of double stranded plasmid DNA representing the
                     library was linearized by NotI. This NotI digested library
                     was used as a template for biotinylated RNA synthesis
                     using SP6 RNA polymerase. Another portion of the double
                     stranded plasmid library was converted to single-stranded
                     circles in vitro using Gene II and Exonuclease III (Life
                     Technologies). Single-stranded DNA (1 mg) was hybridized
                     (C0t 500) with 41 mg of Bio-RNA and vector blocking
                     oligonucleotides. The hybridized Bio-RNA/ss-circles were
                     removed by streptavidin:phenol extraction. EST analysis
                     was performed on the library at the NIH Intramural
                     Sequencing Center(NISC)."
ORIGIN

Alignment Scores:
Pred. No.:              2.77e-87        Length:         642
Score:                  966.00          Matches:        176
Percent Similarity:     96.70%          Conservative:   0
Best Local Similarity:  96.70%          Mismatches:     6
Query Match:            69.35%          Indels:         0
DB:                     14              Gaps:           0

US-09-598-982-21 (1-249) x CD672689 (1-642)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         96 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 155

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        156 CGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 215

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        216 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 275

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        276 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 335

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        336 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 395

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db        396 GTGAACGTCTCCAGCCACGTCCACACGGTCGCCCTGCCCCCTGCCTCAGAGACCTTCCCC 455

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        456 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 515

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        516 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 575

Qy        165 LysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeu 184
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        576 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 635

Qy        185 CysAla 186
              ||||||
Db        636 TGTGCC 641

RESULT 9
BX091892
LOCUS       BX091892                 722 bp    mRNA    linear   EST 23-JAN-2003
DEFINITION  BX091892 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGp998M0584 ; IMAGE:110380, mRNA sequence.
ACCESSION   BX091892
VERSION     BX091892.1  GI:27822632
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 722)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998M0584.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
```

```
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .722
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998M0584 ; IMAGE:110380"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

Alignment Scores:
Pred. No.:              1.34e-86        Length:          722
Score:                  960.00          Matches:         171
Percent Similarity:     100.00%         Conservative:    0
Best Local Similarity:  100.00%         Mismatches:      0
Query Match:            68.92%          Indels:          0
DB:                     13              Gaps:            0

US-09-598-982-21 (1-249) x BX091892 (1-722)

Qy         79 SerArgIleIleValHisProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeu 98
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 60

Qy         99 LeuGluLeuGluGluProValLysValSerSerHisValHisThrValThrLeuProPro 118
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 120

Qy        119 AlaSerGluThrPheProProGlyMetProCysTrpValThrGlyTrpGlyAspValAsp 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 180

Qy        139 AsnAspGluArgLeuProProProPheProLeuLysGlnValLysValProIleMetGlu 158
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 240

Qy        159 AsnHisIleCysAspAlaLysTyrHisLeuGlyAlaTyrThrGlyAspAspValArgIle 178
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 300

Qy        179 ValArgAspAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSer 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 360

Qy        199 GlyGlyProLeuValCysLysValAsnGlyThrTrpLeuGlnAlaGlyValValSerTrp 218
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 420

Qy        219 GlyGluGlyCysAlaGlnProAsnArgProGlyIleTyrThrArgValThrTyrTyrLeu 238
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 480

Qy        239 AspTrpIleHisHisTyrValProLysLysPro 249
```

```
              |||||||||||||||||||||||||||||||||
Db        481 GACTGGATCCACCACTATGTCCCCAAAAAGCCG 513

RESULT 10
BG697540
LOCUS       BG697540                 811 bp    mRNA    linear   EST 07-MAY-2001
DEFINITION  602660760F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4803879 5',
            mRNA sequence.
ACCESSION   BG697540
VERSION     BG697540.1  GI:13963867
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 811)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
            Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10700  row: g  column: 16
            High quality sequence stop: 781.
FEATURES             Location/Qualifiers
     source          1. .811
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4803879"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NCI_CGAP_Skn3"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5kb.  Library constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
ORIGIN

Alignment Scores:
Pred. No.:              3.74e-85        Length:          811
Score:                  946.50          Matches:         187
Percent Similarity:     78.75%          Conservative:    2
Best Local Similarity:  77.92%          Mismatches:      7
Query Match:            67.95%          Indels:          44
DB:                     12              Gaps:            3

US-09-598-982-21 (1-249) x BG697540 (1-811)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        100 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 159

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        160 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 219

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        220 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 279

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        280 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 339

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        340 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 399
```

```
Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        400 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 459

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAsp------------ 140
              ||||||||||||||||||||||||||||||||||||||||||||||||
Db        460 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGGTGGGTCTGGG 519

Qy        140 ------------------------------------------------------------ 140

Db        520 GACAGTGGAGGTGGGGCCAGGGTCTTAGCCACAGCCCAGCCCCTGGGCTCCCTCTGGGCT 579

Qy        141 ---------------------------------GluArgLeuPro--------------- 144
                                               :::   ||||||
Db        580 CCAGGTGGGGGTTGCCCGGCCCCTCCTGAGGCTGACCCTCTTCCCCACTGCAGAGCGCTC 639

Qy        145 Pro-ProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAl 164
              |||       |||||||||||||||||||||||||||||||||||||||||||||||||
Db        640 CCACGCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGG 699

Qy        164 aLysTyrHis-LeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAsp-Met 183
               |||   ||| |||||||||||||||||||||||||||||||||||||||||||::: |||
Db        700 CAAGATACACCTTGGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGAACATG 759

Qy        184 LeuCysAlaGlyAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyPro 201
              |||||||||||||||         ||||||||||||||||||||||||||||||
Db        760 CTGTGTGCGGGAAACCCG---GAGGACTCATGCCAGGGCGACTCCGGAGGGCCC 810

RESULT 11
BQ082929
LOCUS       BQ082929                 502 bp    mRNA    linear   EST 04-APR-2002
DEFINITION  K-EST0144841 S14K402 Homo sapiens cDNA clone S14K402-59-B11 5',
            mRNA sequence.
ACCESSION   BQ082929
VERSION     BQ082929.1  GI:19941018
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 502)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 59  row: B  column: 11
            High quality sequence stop: 502.
FEATURES             Location/Qualifiers
     source          1..502
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S14K402-59-B11"
                     /cell_line="K402"
                     /lab_host="Top10F'"
                     /clone_lib="S14K402"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tobacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including EcoR
                     I site by treatment of T4 RNA ligase and the first strand
                     cDNA was synthesized from oligo dT-selected mRNA by
                     priming with dT-tailed vector. The dT-tailed vector was
                     adjusted to have about 60nt. The cDNA vector was
                     circularized with E. coli DNA ligase after digestion of
                     EcoRI which site is also included in vector. An RNA strand
                     converted to a DNA strand by Okayama-Berg method. The
                     obtained cDNA vectors were used for transformation of
                     competent cells E. coli Top10F' by electroporation method.
                     The cDNA libraries constructed by this method are
                     full-length enriched cDNA library."
ORIGIN

Alignment Scores:
Pred. No.:              2.75e-82     Length:        502
Score:                  915.00       Matches:       166
Percent Similarity:     99.40%       Conservative:  0
Best Local Similarity:  99.40%       Mismatches:    1
Query Match:            65.69%       Indels:        0
DB:                     13           Gaps:          0

US-09-598-982-21 (1-249) x BQ082929 (1-502)

Qy         28 TyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeuThrAlaAla 47
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          2 TACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGCG 61

Qy         48 AlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeuArgGluGln 67
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         62 CACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAG 121

Qy         68 HisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHisProGlnPhe 87
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        122 CACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTC 181

Qy         88 TyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluProValLysVal 107
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        182 TACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTC 241

Qy        108 SerSerHisValHisThrValThrLeuProProAlaSerGluThrPheProProGlyMet 127
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        242 TCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATG 301

Qy        128 ProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuProProProPhe 147
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        302 CCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTT 361

Qy        148 ProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAlaLysTyrHis 167
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        362 CCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCAC 421

Qy        168 LeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeuCysAlaGly 187
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        422 CTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGG 481

Qy        188 AsnThrArgArgAspSerCys 194
              |||||||||||||||||||||
Db        482 AACACCCGGAGGGACTCATGC 502

RESULT 12
BU787772
LOCUS       BU787772                 581 bp    mRNA    linear   EST 11-OCT-2002
DEFINITION  io34c08.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6128247
            5' similar to SW:TRYB_HUMAN P20231 BETA-TRYPTASE PRECURSOR ;, mRNA
            sequence.
ACCESSION   BU787772
VERSION     BU787772.1  GI:23835713
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
```

```
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: io34c08.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 427.
FEATURES             Location/Qualifiers
     source          1..581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6128247"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

Alignment Scores:
Pred. No.:              3.55e-81       Length:         581
Score:                  905.00         Matches:        163
Percent Similarity:     98.79%         Conservative:   0
Best Local Similarity:  98.79%         Mismatches:     2
Query Match:            64.97%         Indels:         0
DB:                     13             Gaps:           0

US-09-598-982-21 (1-249) x BU787772 (1-581)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         86 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 145

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        146 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 205

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        206 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 265

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        266 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 325

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        326 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 385

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        386 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 445

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        446 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 505

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        506 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 565

Qy        165 LysTyrHisLeuGly 169
              |||||||||||||||
Db        566 AAATACCACCTTGGC 580

RESULT 13
BG542089
LOCUS       BG542089                 736 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602571309F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4695458 5',
            mRNA sequence.
ACCESSION   BG542089
VERSION     BG542089.1  GI:13534322
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 736)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1521  row: b  column: 03
            High quality sequence stop: 725.
FEATURES             Location/Qualifiers
     source          1..736
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4695458"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_77"
                     /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.9
                     kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              1.34e-79       Length:         736
Score:                  891.00         Matches:        190
Percent Similarity:     90.52%         Conservative:   1
Best Local Similarity:  90.05%         Mismatches:     11
Query Match:            63.96%         Indels:         11
DB:                     12             Gaps:           1

US-09-598-982-21 (1-249) x BG542089 (1-736)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        115 ATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 174

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        175 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 234

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        235 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 294

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        295 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 354

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        355 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 414

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        415 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 474

Qy        125 Pro-GlyMetProCysTrpValThrGlyTrpGlyAspValAsp-AsnAspGluArgLeuP 144
              ||| ||||||||||||||||||||||||||||||||||||||| ||||||||||||   |
Db        475 CCGGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACCAATGATGAGCGCCT-C 533

Qy        144 roProProPheProLeuLysGlnVal-LysValProIleMetGluAsnHisIleCysAsp 163
              ||      |||   |||||||||||| |||||||||||||||||||||||||||||||||
Db        534 CACGGCATTTCCTTCTGAAGCAGGTTGAAGGTCCCCATAATGGAAAACCACATTTGTGAC 593

Qy        164 AlaLysTyrHisLeuGlyAla-TyrThrGlyAspAsp-ValArgIle------ValArgA 181
              ||||||   ||||||||||||    |||||||||::: |||||||||      |||   |
Db        594 GCAAAATAACACCTTGGGGCCTAACACGGGAGACGAAGGTCCGCATTCGTCCGGTGAACG 653

Qy        181 spAspMetLeuCysAlaGlyAsnThrArgArgAspSerCysGlnGly-AspSer-GlyGl 200
              ||         |||||||||||||||||||||||||||||||||||| |||||| |||||
Db        654 ACAT-TGCTGTTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGGACTCCGGGAGG 712

Qy        200 yProLeuValCysLysVal 206
              |||||||||||||||||||
Db        713 GCCCCTGGTGTGCAAGGTG 731

RESULT 14
CF583026/c
LOCUS       CF583026                 929 bp    mRNA    linear   EST 24-SEP-2003
DEFINITION  AGENCOURT_8786077_updated NIH_MGC_137 Mus musculus cDNA clone
            IMAGE:6432257 3', mRNA sequence.
ACCESSION   CF583026
VERSION     CF583026.1  GI:35196288
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gerard Gradwohl (PNAS 97 P1607-1611, 2000)
             cDNA Library Preparation: Catherine Lee, Endocrine Pancreas
            Consortium
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: IRBD24  row: h  column: 06
            High quality sequence stop: 546.
FEATURES             Location/Qualifiers
     source          1..929
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:6432257"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_137"
                     /note="Organ: pancreas; Vector: pSPORT1; Site_1: SalI;
                     Site_2: NotI; Library consists of a pool of clones
                     rearrayed from the following libraries: Melton normalized
                     mixed mouse pancreas 1 N1-MMS1, Amplified Melton mouse
                     islets 1 MIS1-A, and Kaestner ngn3 wt. Clones rearrayed in
                     the laboratory of K. Kaestner (University of
                     Pennsylvania). Note: this is a NIH_MGC Library."
ORIGIN

Alignment Scores:
Pred. No.:              8.49e-76        Length:         929
Score:                  855.00          Matches:        161
Percent Similarity:     81.00%          Conservative:   18
Best Local Similarity:  72.85%          Mismatches:     35
Query Match:            61.38%          Indels:         8
DB:                     14              Gaps:           3

US-09-598-982-21 (1-249) x CF583026 (1-929)

Qy         32 PheCysGlyGlySerLeuIleHisProGlnTrpValLeuThrAlaAlaAlaCys------ 49
              ||||||   |||         ||||||                  |||   |||
Db        927 TTCTGCCGAGGG---TTTTTCATCCCACC---CCCAAGGGGGGTGCTCANTGCGGGCAC 874

Qy         50 ------ValGlyProAspValLysAspLeuAlaAlaLeu-ArgValGlnLeuArgGluGl 67
                    |||   |||   :::|||         |||||| ||||||||||||||||||||
Db        873 ACTGGTGTGGACCCGCACATCAAAAAGCCCACAGCTCTTCCGGGTGCAGCTTCGTGAGCA 814

Qy         67 nHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHisProGlnPh 87
              |:::|||||||||   ||||||   |||   ::::::||||||:::||||||||    ::
Db        813 GTATCTATACTATGGGACCAGCT-CTCTCTTTGAACCGGATCGTGGTGCACCCCCACTA 755

Qy         87 eTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluProValLysVa 107
              :|||||||||:::   |||||||||:::||||||||||||||||||   ||||||   ||
Db        754 TTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCCCTGTGAATGT 695

Qy        107 lSerSerHisValHisThrValThrLeuProProAlaSerGluThrPheProProGlyMe 127
              ||||:::|||:::|||   ::::::|||||||||||||||||||||||||||||||||
Db        694 CTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCCCCCCTGGGAC 635

Qy        127 tProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuProProProPh 147
                  ||||||||||||||||||||||||:::||||||||||||   ||||||||||||::
Db        634 ATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCCCACCTCCTTA 575

Qy        147 eProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAlaLysTyrHi 167
              :|||||||||||||||||||||||||||:::||||||   :::||||||   ||||||||
Db        574 TCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACCGGAAGTACCA 515

Qy        167 sLeuGlyAlaTyrThrGlyAspAspValArgIleValArgAspAspMetLeuCysAlaGl 187
              |   |||   |||||||||||||||      ||||||   |||   ||||||||||||||
Db        514 CACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGCTGTGTGCTGG 455

Qy        187 yAsnThrArgArgAspSerCysGlnGlyAspSerGlyGlyProLeuValCysLysValAs 207
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 AAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCTGCAAAGTGAA 395

Qy        207 nGlyThrTrpLeuGlnAlaGlyValValSerTrpGlyGluGlyCysAlaGlnProAsnAr 227
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||::
Db        394 GGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGGTGAGGGCTGCGCACAGCCCAACAA 335

Qy        227 gProGlyIleTyrThrArgValThrTyrTyrLeuAspTrpIleHisHisTyrValProLy 247
              :|||||||||||||||||||||||||||||||||||||||||||||   |||||||||::
Db        334 GCCTGGCATCTACACCCGGGTGACATACTACTTAGACTGGATCCACCGCTATGTCCCTGA 275

Qy        247 s 247
```

```
              :
Db        274 G 274

RESULT 15
CF553058
LOCUS       CF553058                 806 bp    mRNA    linear   EST 22-SEP-2003
DEFINITION  AGENCOURT_15594982 NIH_MGC_183 Homo sapiens cDNA clone
            IMAGE:30529468 5', mRNA sequence.
ACCESSION   CF553058
VERSION     CF553058.1  GI:34889892
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 806)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM617  row: b  column: 05
            High quality sequence start: 18
            High quality sequence stop: 611.
FEATURES             Location/Qualifiers
     source          1..806
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30529468"
                     /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
                     /clone_lib="NIH_MGC_183"
                     /note="Organ: Pooled muscle (cardiac and skeletal);
                     Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                     NotI; Library is oligo-dT primed and directionally cloned
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 1.7. Library was constructed by Invitrogen."
ORIGIN

Alignment Scores:
Pred. No.:              6.24e-75      Length:        806
Score:                  845.50        Matches:       159
Percent Similarity:     89.01%        Conservative:  3
Best Local Similarity:  87.36%        Mismatches:    17
Query Match:            60.70%        Indels:        3
DB:                     14            Gaps:          1

US-09-598-982-21 (1-249) x CF553058 (1-806)

Qy          5 IleValGlyGlyGlnGluAlaProArgSerLysTrpProTrpGlnValSerLeuArgVal 24
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        248 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 307

Qy         25 HisGlyProTyrTrpMetHisPheCysGlyGlySerLeuIleHisProGlnTrpValLeu 44
                       |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        308 CGCGACCGATACTGGATGCACTTCTGTGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 367

Qy         45 ThrAlaAlaAlaCysValGlyProAspValLysAspLeuAlaAlaLeuArgValGlnLeu 64
              |||||||||   |||:::||||||||||||||||||||||||   |||||||||||||||
Db        368 ACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTG 427

Qy         65 ArgGluGlnHisLeuTyrTyrGlnAspGlnLeuLeuProValSerArgIleIleValHis 84
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        428 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 487

Qy         85 ProGlnPheTyrThrAlaGlnIleGlyAlaAspIleAlaLeuLeuGluLeuGluGluPro 104
              ||||||||||||      |||   ||||||||||||||||||||||||||||||||||||
Db        488 CCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCC 547

Qy        105 ValLysValSerSerHisValHisThrValThrLeuProProAlaSerGluThrPhePro 124
              |||   :::||||||   ||||||||||||   |||||||||||||||||||||||||||
Db        548 GTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCC 607

Qy        125 ProGlyMetProCysTrpValThrGlyTrpGlyAspValAspAsnAspGluArgLeuPro 144
              |||||||||||||||||||||||||||||||||||||||||||||||||||   ||||||
Db        608 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCA 667

Qy        145 ProProPheProLeuLysGlnValLysValProIleMetGluAsnHisIleCysAspAla 164
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        668 CCGCCATTTCCCCTGAAGCAGGTGAAAGTCCCCATAATGGAAAACCACATTTGTGACGCA 727

Qy        165 LysTyrHisLeuGly--AlaTyrThrGlyAspAsp---ValArgIleValArgAspAspM 183
              |||||||||||||||     ||||||||||||:::               ||||||||||
Db        728 AAATACCACCTTGGGCGCCCTACACGGGAGACGAACGTCCGGCATCATCCCGTGACGACA 787

Qy        183 et 183
              ||
Db        788 TG 789

Search completed: July 22, 2004, 18:14:21
Job time : 3075 secs
```

```
                 GenCore version 5.1.6
        Copyright (c) 1993 - 2004  Compugen Ltd.


OM nucleic - nucleic search, using sw model

Run on:          July 22, 2004, 10:05:00 ; Search time 5131 Seconds
                                            (without alignments)
                                            6512.859 Million cell updates/sec

Title:           US-09-598-982-20
Perfect score:   771
Sequence:        1 gggcccctcgagaaaagaat..........cgtgaagcggccgccgtcgt 771

Scoring table:   IDENTITY_NUC
                 Gapop 10.0 , Gapext 1.0

Searched:        3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters:      6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :        GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*
                 15:  em_ba:*
                 16:  em_fun:*
                 17:  em_hum:*
                 18:  em_in:*
                 19:  em_mu:*
                 20:  em_om:*
                 21:  em_or:*
                 22:  em_ov:*
                 23:  em_pat:*
                 24:  em_ph:*
                 25:  em_pl:*
                 26:  em_ro:*
                 27:  em_sts:*
                 28:  em_un:*
                 29:  em_vi:*
                 30:  em_htg_hum:*
                 31:  em_htg_inv:*
                 32:  em_htg_other:*
                 33:  em_htg_mus:*
                 34:  em_htg_pln:*
                 35:  em_htg_rod:*
                 36:  em_htg_mam:*
                 37:  em_htg_vrt:*
                 38:  em_sy:*
                 39:  em_htgo_hum:*
                 40:  em_htgo_mus:*
                 41:  em_htgo_other:*

      Pred. No. is the number of results predicted by chance to have a
```

```
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                    SUMMARIES
                   %
Result           Query
  No.   Score    Match Length DB  ID                             Description
 ---------------------------------------------------------------------------------
    1     771    100.0    771  6  AX347853                       AX347853 Sequence
    2   769.4     99.8    771  6  AX347869                       AX347869 Sequence
    3   764.6     99.2    771  6  AR165112                       AR165112 Sequence
    4   764.6     99.2    771  6  AX347841                       AX347841 Sequence
    5   764.6     99.2    771  6  BD211490                       BD211490 Enzymatic
    6     763     99.0    771  6  AX347838                       AX347838 Sequence
    7   761.4     98.8    771  6  AX347855                       AX347855 Sequence
    8   759.8     98.5    771  6  AX347871                       AX347871 Sequence
    9     755     97.9    771  6  AX347857                       AX347857 Sequence
   10     755     97.9    771  6  AX347859                       AX347859 Sequence
   11   753.4     97.7    771  6  AX347873                       AX347873 Sequence
   12   753.4     97.7    771  6  AX347875                       AX347875 Sequence
   13     735     95.3    735  6  AX347861                       AX347861 Sequence
   14   733.4     95.1    735  6  AX347877                       AX347877 Sequence
   15     731     94.8   1128  6  AR080461                       AR080461 Sequence
   16     731     94.8   1128  9  HUMTRY2A                       M33492 Human trypt
   17     731     94.8   1137  6  AR080460                       AR080460 Sequence
   18     731     94.8   1137  9  HUMTRY1A                       M33491 Human trypt
   19     731     94.8   1143  9  HUMBTRYP                       M37488 Human beta-
   20     731     94.8   1145  9  S55551                         S55551 beta-trypta
   21     731     94.8   1194  9  BC029356                       BC029356 Homo sapi
   22     731     94.8   1235  9  BC038416                       BC038416 Homo sapi
   23   728.6     94.5    735  6  AR165109                       AR165109 Sequence
   24   728.6     94.5    735  6  AX347843                       AX347843 Sequence
   25   728.6     94.5    735  6  BD211487                       BD211487 Enzymatic
   26     727     94.3    735  6  AX347834                       AX347834 Sequence
   27   726.2     94.2   1081  6  AR080462                       AR080462 Sequence
   28   726.2     94.2   1081  6  AX329646                       AX329646 Sequence
   29   726.2     94.2   1081  6  AX330205                       AX330205 Sequence
   30   726.2     94.2   1081  6  AX334644                       AX334644 Sequence
   31   726.2     94.2   1081  6  AX375047                       AX375047 Sequence
   32   726.2     94.2   1081  9  HUMTRY3A                       M33493 Human trypt
   33   725.4     94.1    735  6  AX347863                       AX347863 Sequence
   34   723.8     93.9    735  6  AX347879                       AX347879 Sequence
   35     719     93.3    735  6  AX347865                       AX347865 Sequence
   36     719     93.3    735  6  AX347867                       AX347867 Sequence
   37     719     93.3   1221  6  AX014344                       AX014344 Sequence
   38     719     93.3   1221  6  BD222223                       BD222223 Human nuc
   39   718.2     93.2   2662  6  AX338508                       AX338508 Sequence
   40   717.4     93.0    735  6  AX347881                       AX347881 Sequence
   41   717.4     93.0    735  6  AX347883                       AX347883 Sequence
   42   692.6     89.8    884  9  AF206665                       AF206665 Homo sapi
   43   692.6     89.8   1206  9  BC028059                       BC028059 Homo sapi
   44   692.6     89.8   1207  9  BC051852                       BC051852 Homo sapi
   45   686.2     89.0   1154  6  AR080459                       AR080459 Sequence


                                    ALIGNMENTS


RESULT 1
AX347853
LOCUS       AX347853                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 20 from Patent WO0198470.
ACCESSION   AX347853
VERSION     AX347853.1  GI:18495660
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
```

```
  JOURNAL   Patent: WO 0198470-A 20 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <1..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22420.1"
                     /db_xref="GI:18495661"
                     /db_xref="REMTREMBL:CAD22420"
                     /translation="GPLEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQW
                     VLTAAACVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLE
                     LEEPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIME
                     NHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVV
                     SWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match            100.0%;  Score 771;  DB 6;  Length 771;
  Best Local Similarity  100.0%;  Pred. No. 2.3e-117;
  Matches  771;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy             1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy            61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy           121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy           181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy           241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy           301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy           361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy           421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy           481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy           541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy           601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy           661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy           721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
```

```
                 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db           721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 2
AX347869
LOCUS       AX347869                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 36 from Patent WO0198470.
ACCESSION   AX347869
VERSION     AX347869.1  GI:18495676
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 36 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22428.1"
                     /db_xref="GI:18495677"
                     /db_xref="REMTREMBL:CAD22428"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAACVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match            99.8%;  Score 769.4;  DB 6;  Length 771;
  Best Local Similarity  99.9%;  Pred. No. 4.2e-117;
  Matches  770;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy             1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy            61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy           121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy           181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy           241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy           301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
                 ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db           301 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy           361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy           421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db           421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy           481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy           541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy           601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy           661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy           721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
                 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db           721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 3
AR165112
LOCUS       AR165112                 771 bp    DNA     linear   PAT 17-OCT-2001
DEFINITION  Sequence 4 from patent US 6274366.
ACCESSION   AR165112
VERSION     AR165112.1  GI:16238527
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 771)
  AUTHORS   Maffitt,M.A., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Enzymatically-active recombinant human .beta.-tryptase and method
            of making same
  JOURNAL   Patent: US 6274366-A 4 14-AUG-2001;
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             99.2%;  Score 764.6;  DB 6;  Length 771;
  Best Local Similarity   99.5%;  Pred. No. 2.6e-116;
  Matches  767;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy             1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy            61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy           121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCC 180
                 ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db           121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCC 180

Qy           181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy           241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy           301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy           361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy           421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy           481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy           541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy           601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy           661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy           721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
                 |||||||||||||||||||||||||||||||||||||||||||||||||||
Db           721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 4
AX347841
LOCUS       AX347841                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 8 from Patent WO0198470.
ACCESSION   AX347841
VERSION     AX347841.1  GI:18495648
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 8 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7. .756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22418.1"
                     /db_xref="GI:18495649"
                     /db_xref="REMTREMBL:CAD22418"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match             99.2%;  Score 764.6;  DB 6;  Length 771;
  Best Local Similarity   99.5%;  Pred. No. 2.6e-116;
  Matches  767;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy             1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db             1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy            61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db           61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy          121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
                ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db          121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy          181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy          301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy          661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy          721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771

RESULT 5
BD211490
LOCUS       BD211490                 771 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Enzymatically-active recombinant human beta-tryptase and method of
            making same.
ACCESSION   BD211490
VERSION     BD211490.1  GI:33021260
KEYWORDS    JP 2002515254-A/4.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 771)
  AUTHORS   Maffitt,M.A., Niles,A.L. and Frendscho,M.H.
  TITLE     Enzymatically-active recombinant human beta-tryptase and method of
            making same
  JOURNAL   Patent: JP 2002515254-A 4 28-MAY-2002;
            PROMEGA CORP
COMMENT     OS   Homo sapiens (human)
            PN   JP 2002515254-A/4
            PD   28-MAY-2002
            PF   29-OCT-1998 JP 2000549745
            PR   15-MAY-1998 US     09/079970
            PI   MARK A MAFFITT,ANDREW L NILES,MARY HAAK FRENDSCHO PC
            C12N15/09,C07K16/40,C12N1/19,C12N9/64,C12P21/08,C12Q1/37// PC
            (C12N9/64,C12R1:645),C12N15/00
```

```
            CC   Strandedness: Single;
            CC   Topology: Linear;
            CC   Enzymatically-active recombinant human beta-tryptase and  CC
                    method of making
            CC   same
            FH   Key             Location/Qualifiers
            FT   CDS             7. .753
            FT   misc_signal     7. .18.
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
ORIGIN

  Query Match             99.2%;  Score 764.6;  DB 6;  Length 771;
  Best Local Similarity   99.5%;  Pred. No. 2.6e-116;
  Matches  767;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy            1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy           61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy          121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
                ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db          121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy          181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy          301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy          661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy          721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771

RESULT 6
```

```
AX347838
LOCUS       AX347838                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 5 from Patent WO0198470.
ACCESSION   AX347838
VERSION     AX347838.1  GI:18495645
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 5 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22417.1"
                     /db_xref="GI:18495646"
                     /db_xref="REMTREMBL:CAD22417"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match             99.0%;  Score 763;  DB 6;  Length 771;
  Best Local Similarity   99.4%;  Pred. No. 4.8e-116;
  Matches  766;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 7
AX347855
LOCUS       AX347855                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 22 from Patent WO0198470.
ACCESSION   AX347855
VERSION     AX347855.1  GI:18495662
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 22 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22421.1"
                     /db_xref="GI:18495663"
                     /db_xref="REMTREMBL:CAD22421"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVQPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAAIALLELE
                     EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match             98.8%;  Score 761.4;  DB 6;  Length 771;
  Best Local Similarity   99.2%;  Pred. No. 8.7e-116;
  Matches  765;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
```

```
Qy    241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db    241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTG 300

Qy    301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy    361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy    421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy    481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy    541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy    601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy    661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy    721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771

RESULT 8
AX347871
LOCUS       AX347871                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 38 from Patent WO0198470.
ACCESSION   AX347871
VERSION     AX347871.1  GI:18495678
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 38 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22429.1"
                     /db_xref="GI:18495679"
                     /db_xref="REMTREMBL:CAD22429"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGAAIALLELE
                     EPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match             98.5%;  Score 759.8;  DB 6;  Length 771;
  Best Local Similarity   99.1%;  Pred. No. 1.6e-115;
  Matches  764;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy      1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy     61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy    121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
          ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db    121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy    181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy    241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
          |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db    241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTG 300

Qy    301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
          ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db    301 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy    361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy    421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy    481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy    541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy    601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy    661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy    721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771

RESULT 9
AX347857
LOCUS       AX347857                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 24 from Patent WO0198470.
ACCESSION   AX347857
VERSION     AX347857.1  GI:18495664
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 24 27-DEC-2001;
            PROMEGA CORPORATION (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7. .756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22422.1"
                     /db_xref="GI:18495665"
                     /db_xref="REMTREMBL:CAD22422"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match             97.9%;  Score 755;  DB 6;  Length 771;
  Best Local Similarity   98.7%;  Pred. No. 9.9e-115;
  Matches  761;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||| || |||||| ||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              || || || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCGGACCTCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
```

```
RESULT 10
AX347859
LOCUS       AX347859                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 26 from Patent WO0198470.
ACCESSION   AX347859
VERSION     AX347859.1  GI:18495666
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 26 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7. .756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22423.1"
                     /db_xref="GI:18495667"
                     /db_xref="REMTREMBL:CAD22423"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVKVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match             97.9%;  Score 755;  DB 6;  Length 771;
  Best Local Similarity   98.7%;  Pred. No. 9.9e-115;
  Matches  761;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
```

```
Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              |||||||||||||||||||||||||||||||||||||||||||||||||| || ||| ||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              || || || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCGGACCACTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771

RESULT 11
AX347873
LOCUS       AX347873                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 40 from Patent WO0198470.
ACCESSION   AX347873
VERSION     AX347873.1  GI:18495680
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 40 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <1..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22430.1"
                     /db_xref="GI:18495681"
                     /db_xref="REMTREMBL:CAD22430"
                     /translation="GPLEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQW
                     VLTAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLE
                     LEEPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIME
                     NHICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVV
                     SWGEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match             97.7%;  Score 753.4;  DB 6;  Length 771;
  Best Local Similarity   98.6%;  Pred. No. 1.8e-114;
  Matches  760;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||| || |||||| ||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              || || || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCGGACCTCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771

RESULT 12
AX347875
LOCUS       AX347875                 771 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 42 from Patent WO0198470.
ACCESSION   AX347875
VERSION     AX347875.1  GI:18495682
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 42 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..771
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <7..756
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22431.1"
                     /db_xref="GI:18495683"
                     /db_xref="REMTREMBL:CAD22431"
                     /translation="LEKRIVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVL
                     TAAHCVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELE
                     EPVNVSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENH
                     ICDAKYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDAGGPLVCKVNGTWLQAGVVSW
                     GEGCAQPNRPGIYTRVTYYLDWIHHYVPKKP"
```

```
ORIGIN

  Query Match             97.7%;  Score 753.4;  DB 6;  Length 771;
  Best Local Similarity   98.6%;  Pred. No. 1.8e-114;
  Matches  760;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              |||||||||||||||||||||||||||||||||||||||||||||||||| || ||| ||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              || || || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCGGACCACTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 13
AX347861
LOCUS       AX347861                 735 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 28 from Patent WO0198470.
ACCESSION   AX347861
VERSION     AX347861.1  GI:18495668
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 28 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1. .735
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <1. .>735
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22424.1"
                     /db_xref="GI:18495669"
                     /db_xref="REMTREMBL:CAD22424"
                     /translation="IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAA
                     CVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVK
                     VSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDA
                     KYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGC
                     AQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match             95.3%;  Score 735;  DB 6;  Length 735;
  Best Local Similarity   100.0%;  Pred. No. 2e-111;
  Matches  735;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy         79 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy        139 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy        199 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy        259 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy        319 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy        379 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy        439 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy        499 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy        559 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy        619 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy        679 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720
```

```
Qy        739 GTCCCCAAAAAGCCG 753
              |||||||||||||||
Db        721 GTCCCCAAAAAGCCG 735

RESULT 14
AX347877
LOCUS       AX347877                 735 bp    DNA     linear   PAT 01-FEB-2002
DEFINITION  Sequence 44 from Patent WO0198470.
ACCESSION   AX347877
VERSION     AX347877.1  GI:18495684
KEYWORDS    .
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Maffit,M., Niles,A.L. and Haak-Frendscho,M.
  TITLE     Recombinant proteolytic tryptases, active site mutants thereof, and
            methods of making same
  JOURNAL   Patent: WO 0198470-A 44 27-DEC-2001;
            PROMEGA CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..735
                     /organism="Homo sapiens"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:9606"
     CDS             <1..>735
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD22432.1"
                     /db_xref="GI:18495685"
                     /db_xref="REMTREMBL:CAD22432"
                     /translation="IVGGQEAPRSKWPWQVSLRVHGPYWMHFCGGSLIHPQWVLTAAA
                     CVGPDVKDLAALRVQLREQHLYYQDQLLPVSRIIVHPQFYTAQIGADIALLELEEPVN
                     VSSHVHTVTLPPASETFPPGMPCWVTGWGDVDNDERLPPPFPLKQVKVPIMENHICDA
                     KYHLGAYTGDDVRIVRDDMLCAGNTRRDSCQGDSGGPLVCKVNGTWLQAGVVSWGEGC
                     AQPNRPGIYTRVTYYLDWIHHYVPKKP"
ORIGIN

  Query Match             95.1%;  Score 733.4;  DB 6;  Length 735;
  Best Local Similarity   99.9%;  Pred. No. 3.6e-111;
  Matches  734;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy         19 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy         79 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy        139 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy        199 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy        259 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy        319 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 378
              ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy        379 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy        439 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy        499 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy        559 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy        619 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy        679 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy        739 GTCCCCAAAAAGCCG 753
              |||||||||||||||
Db        721 GTCCCCAAAAAGCCG 735

RESULT 15
AR080461
LOCUS       AR080461                1128 bp    DNA     linear   PAT 31-AUG-2000
DEFINITION  Sequence 15 from patent US 5968782.
ACCESSION   AR080461
VERSION     AR080461.1  GI:10007196
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1128)
  AUTHORS   Stevens,R.L.
  TITLE     Mast cell protease that cleaves fibrinogen
  JOURNAL   Patent: US 5968782-A 15 19-OCT-1999;
FEATURES             Location/Qualifiers
     source          1..1128
                     /organism="unknown"
                     /mol_type="unassigned DNA"
ORIGIN

  Query Match             94.8%;  Score 731;  DB 6;  Length 1128;
  Best Local Similarity   98.7%;  Pred. No. 8.3e-111;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 146

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 206

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 266

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 326

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 386

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 446

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          447 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 506

Qy          437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          507 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 566

Qy          497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          567 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 626

Qy          557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          627 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 686

Qy          617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          687 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 746

Qy          677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          747 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 806

Qy          737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
                ||||||||||||||||||||  | | |
Db          807 ATGTCCCCAAAAAGCCGTGAGTCAGGC 833

Search completed: July 22, 2004, 12:27:24
Job time : 5134 secs
```

```
                   GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         July 22, 2004, 12:27:31 ; Search time 569 Seconds
                                           (without alignments)
                                           6623.689 Million cell updates/sec

Title:          US-09-598-982-20
Perfect score:  771
Sequence:       1 gggcccctcgagaaaagaat..........cgtgaagcggccgccgtcgt 771

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       3216467 seqs, 2444149694 residues

Total number of hits satisfying chosen parameters:      6432934

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Published_Applications_NA:*
                 1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                 2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                 3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                 4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                 5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                 6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                 7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                 8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                 9:  /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
                 10:  /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
                 11:  /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
                 12:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                 13:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
                 14:  /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
                 15:  /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
                 16:  /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
                 17:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                 18:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                 19:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

       Pred. No. is the number of results predicted by chance to have a
       score greater than or equal to the score of the result being printed,
       and is derived by analysis of the total score distribution.

                                    SUMMARIES
                    %
Result            Query
  No.    Score    Match Length DB  ID                          Description
 ------------------------------------------------------------------------------
     1      731    94.8   1143  16  US-10-352-684A-47              Sequence 47, Appl
     2      731    94.8   1145  17  US-10-287-226-93               Sequence 93, Appl
     3    730.4    94.7    828  17  US-10-287-226-91               Sequence 91, Appl
     4    726.2    94.2   1081   9  US-09-954-456-2126           Sequence 2126, Ap
     5    726.2    94.2   1081  10  US-09-960-706-680              Sequence 680, App
     6    726.2    94.2   1081  10  US-09-873-319-427              Sequence 427, App
     7    726.2    94.2   1081  10  US-09-873-367C-155             Sequence 155, App
     8    726.2    94.2   1081  10  US-09-873-367C-714             Sequence 714, App
     9    718.2    93.2   2662  17  US-10-275-505-27               Sequence 27, Appl
    10      700    90.8    828  17  US-10-287-226-89               Sequence 89, Appl
    11    686.2    89.0   1154  16  US-10-352-684A-45              Sequence 45, Appl
    12    686.2    89.0   1158  13  US-10-116-802-240              Sequence 240, App
    13      678    87.9    858  17  US-10-287-226-87               Sequence 87, Appl
    14      463    60.1   1187  17  US-10-275-505-22               Sequence 22, Appl
```

```
     15    287.4    37.3    846  10  US-09-813-432-11              Sequence 11, Appl
     16    287.4    37.3    846  13  US-10-246-583-11              Sequence 11, Appl
     17    287.4    37.3    846  16  US-10-174-364-11              Sequence 11, Appl
     18    287.4    37.3    846  17  US-10-689-832-11              Sequence 11, Appl
     19      275    35.7    383  10  US-09-918-995-6378            Sequence 6378, Ap
     20    252.4    32.7    948  13  US-10-150-813-15              Sequence 15, Appl
     21    252.4    32.7    948  13  US-10-139-854-15              Sequence 15, Appl
     22    252.4    32.7    948  15  US-10-131-409-15              Sequence 15, Appl
     23    252.4    32.7    948  16  US-10-150-811-15              Sequence 15, Appl
     24    236.4    30.7    858  13  US-10-150-813-21              Sequence 21, Appl
     25    236.4    30.7    858  13  US-10-139-854-21              Sequence 21, Appl
     26    236.4    30.7    858  15  US-10-131-409-21              Sequence 21, Appl
     27    236.4    30.7    858  16  US-10-150-811-21              Sequence 21, Appl
     28    233.4    30.3    873  15  US-10-117-323-2               Sequence 2, Appli
     29    231.6    30.0    843  13  US-10-150-813-25              Sequence 25, Appl
     30    231.6    30.0    843  13  US-10-139-854-25              Sequence 25, Appl
     31    231.6    30.0    843  15  US-10-131-409-25              Sequence 25, Appl
     32    231.6    30.0    843  16  US-10-150-811-25              Sequence 25, Appl
     33      230    29.8    978  15  US-10-311-955-1               Sequence 1, Appli
 c   34    223.2    28.9    616   9  US-09-954-456-2242          Sequence 2242, Ap
 c   35    223.2    28.9    616  10  US-09-960-706-74              Sequence 74, Appl
 c   36    223.2    28.9    616  10  US-09-873-319-51              Sequence 51, Appl
 c   37    223.2    28.9    616  10  US-09-873-367C-545            Sequence 545, App
     38    221.8    28.8    870  13  US-10-150-813-19              Sequence 19, Appl
     39    221.8    28.8    870  13  US-10-139-854-19              Sequence 19, Appl
     40    221.8    28.8    870  15  US-10-131-409-19              Sequence 19, Appl
     41    221.8    28.8    870  16  US-10-150-811-19              Sequence 19, Appl
     42    218.8    28.4   1958  17  US-10-311-035-29              Sequence 29, Appl
     43    201.4    26.1    780  15  US-10-266-035-3               Sequence 3, Appli
     44    201.4    26.1    783  17  US-10-391-364-22              Sequence 22, Appl
     45    201.4    26.1    980  15  US-10-266-035-1               Sequence 1, Appli

                                   ALIGNMENTS


RESULT 1
US-10-352-684A-47
; Sequence 47, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
```

```
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 47
;   LENGTH: 1143
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (6)...(833)
US-10-352-684A-47

  Query Match             94.8%;  Score 731;  DB 16;  Length 1143;
  Best Local Similarity   98.0%;  Pred. No. 4.4e-193;
  Matches  740;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         94 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 153

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        154 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 213

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        214 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 273

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        274 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 333

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        334 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 393

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        394 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 453

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        454 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 513

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        514 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 573

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        574 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 633

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        634 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 693

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        694 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 753

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        754 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 813

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              ||||||||||||||||||||  | | |   |  ||
Db        814 ATGTCCCCAAAAAGCCGTGAGTCAGGCCTGGGTGT 848

RESULT 2
US-10-287-226-93
; Sequence 93, Application US/10287226
```

```
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 93
;   LENGTH: 1145
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (8)..(832)
```

```
US-10-287-226-93

  Query Match             94.8%;  Score 731;  DB 17;  Length 1145;
  Best Local Similarity   98.7%;  Pred. No. 4.4e-193;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy           17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
                | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           96 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 155

Qy           77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          156 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 215

Qy          137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
                ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db          216 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 275

Qy          197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          276 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 335

Qy          257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          336 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 395

Qy          317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          396 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 455

Qy          377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          456 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 515

Qy          437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          516 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 575

Qy          497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          576 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 635

Qy          557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          636 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 695

Qy          617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          696 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 755

Qy          677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          756 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 815

Qy          737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
                ||||||||||||||||||||  | | |
Db          816 ATGTCCCCAAAAAGCCGTGAGTCAGGC 842


RESULT 3
US-10-287-226-91
; Sequence 91, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 91
;   LENGTH: 828
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(825)
US-10-287-226-91

  Query Match             94.7%;  Score 730.4;  DB 17;  Length 828;
  Best Local Similarity   99.2%;  Pred. No. 6.1e-193;
  Matches  734;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy           17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
                | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           89 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 148

Qy           77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 208

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        209 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 268

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 328

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        329 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 388

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        389 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 448

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        449 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 508

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 568

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 628

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 688

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 748

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 808

Qy        737 ATGTCCCCAAAAAGCCGTGA 756
              ||||||||||||||||||||
Db        809 ATGTCCCCAAAAAGCCGTGA 828

RESULT 4
US-09-954-456-2126
; Sequence 2126, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
;  TITLE OF INVENTION:  Sets
;  FILE REFERENCE: 689290-76
;  CURRENT APPLICATION NUMBER: US/09/954,456
;  CURRENT FILING DATE:  2001-09-18
;  PRIOR APPLICATION NUMBER: US/60/233,617
;  PRIOR FILING DATE: 2000-09-18
;  PRIOR APPLICATION NUMBER: US/60/234,052
;  PRIOR FILING DATE: 2000-09-20
;  PRIOR APPLICATION NUMBER: US/60/234,923
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,134
;  PRIOR FILING DATE: 2000-09-25
;  PRIOR APPLICATION NUMBER: US/60/235,637
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,638
;  PRIOR FILING DATE: 2000-09-26
;  PRIOR APPLICATION NUMBER: US/60/235,711
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,720
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,840
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: US/60/235,863
;  PRIOR FILING DATE: 2000-09-27
;  NUMBER OF SEQ ID NOS: 2276
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2126
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-954-456-2126

  Query Match            94.2%;  Score 726.2;  DB 9;  Length 1081;
  Best Local Similarity  98.3%;  Pred. No. 9.3e-192;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812

RESULT 5
US-09-960-706-680
```

```
; Sequence 680, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplas
;  TITLE OF INVENTION:  Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-01US
;  CURRENT APPLICATION NUMBER: US/09/960,706
;  CURRENT FILING DATE:  2001-09-24
;  PRIOR APPLICATION NUMBER: 60/223,323
;  PRIOR FILING DATE: 2000-08-07
;  PRIOR APPLICATION NUMBER: 09/873,319
;  PRIOR FILING DATE: 2001-06-05
;  NUMBER OF SEQ ID NOS: 1124
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 680
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134280A1 M33493
US-09-960-706-680

  Query Match             94.2%;  Score 726.2;  DB 10;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 9.3e-192;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812


RESULT 6
US-09-873-319-427
; Sequence 427, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
;  APPLICANT: Munger, William E.
;  APPLICANT: Kulkarni, Prakash
;  APPLICANT: Getzenberg, Robert H.
;  APPLICANT: Waga, Iwao
;  APPLICANT: Yamamoto, Jun
;  TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic
;  TITLE OF INVENTION: Hyperplasia Using Gene Expression Profiles
;  FILE REFERENCE: 44921-5029-US
;  CURRENT APPLICATION NUMBER: US/09/873,319A
;  CURRENT FILING DATE: 2001-06-05
;  EARLIER APPLICATION NUMBER: US 60/223,323
;  EARLIER FILING DATE: 2000-08-07
;  NUMBER OF SEQ ID NOS: 755
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 427
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20030134324A1 M33493
US-09-873-319-427

  Query Match             94.2%;  Score 726.2;  DB 10;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 9.3e-192;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605
```

```
Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812

RESULT 7
US-09-873-367C-155
; Sequence 155, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 155
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-155

  Query Match             94.2%;  Score 726.2;  DB 10;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 9.3e-192;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365
```

```
Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812

RESULT 8
US-09-873-367C-714
; Sequence 714, Application US/09873367C
; Publication No. US20030165839A1
; GENERAL INFORMATION:
;  APPLICANT: Young, Paul
;  APPLICANT:  Soppet, Daniel
;  APPLICANT:  Endress, Gregory
;  APPLICANT:  Augustus, Meena
;  APPLICANT:  Ebner, Reinhard
;  APPLICANT:  Carter, Kenneth
;  TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using
;  TITLE OF INVENTION:  Signature Gene Sets
;  FILE REFERENCE: 689290-64
;  CURRENT APPLICATION NUMBER: US/09/873,367C
;  CURRENT FILING DATE:  2003-04-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,891
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/236,842
;  PRIOR FILING DATE: 2000-09-29
;  PRIOR APPLICATION NUMBER: U.S. 60/244,867
;  PRIOR FILING DATE: 2000-11-01
;  PRIOR APPLICATION NUMBER: U.S. 60/245,084
;  PRIOR FILING DATE: 2000-11-01
;  NUMBER OF SEQ ID NOS: 1067
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 714
;   LENGTH: 1081
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-367C-714

  Query Match             94.2%;  Score 726.2;  DB 10;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 9.3e-192;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125
```

```
Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812


RESULT 9
US-10-275-505-27
; Sequence 27, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  DELEGEANE, Angelo M.; LAL, Preeti G.
;  APPLICANT:  HAFALIA, April J.A.; PATTERSON, Chandra
;  APPLICANT:  WALIA, Narinder K.; KEARNEY, Liam
;  APPLICANT:  TRIBOULEY, Catherine M.; KHAN, Farrah A.
;  APPLICANT:  YAO, Monique G.; BAUGHN, Mariah R.
;  APPLICANT:  AZIMZAI, Yalda; ELLIOTT, Vicki S.
;  APPLICANT:  NGUYEN, Danniel B.; GANDHI, Ameena R.
;  APPLICANT:  YANG, Junming; HERNANDEZ, Roberto
;  APPLICANT:  POLICKY, Jennifer L.; LU, Dyung Aina M.
;  APPLICANT:  REDDY, Roopa M.; YUE, Henry
;  APPLICANT:  TANG, Y. Tom
;  TITLE OF INVENTION: PROTEASES
;  FILE REFERENCE: PI-0085 USN
;  CURRENT APPLICATION NUMBER: US/10/275,505
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: PCT/US01/14651
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: 60/209,402
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: 60/207,477
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/205,803
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/203,566
;  PRIOR FILING DATE: 2000-05-11
;  PRIOR APPLICATION NUMBER: 60/202,082
;  PRIOR FILING DATE: 2000-05-04
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PERL Program
; SEQ ID NO 27
;   LENGTH: 2662
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 7472460CB1
US-10-275-505-27

  Query Match             93.2%;  Score 718.2;  DB 17;  Length 2662;
  Best Local Similarity   98.2%;  Pred. No. 1.7e-189;
  Matches  726;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        699 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 758

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        759 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 818

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        819 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 878

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        879 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 938

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        939 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 998

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        999 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 1058

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1059 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 1118

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              |||||||||||||||||||||| |||||||||||||||||||||||||||||||||||||
Db       1119 CACCGCCATTTCCTCTGAAGCATGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 1178

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1179 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 1238

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1239 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 1298

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              |||||||||||||||||||||||||||||||||||||||| ||||| |||||||||||||
Db       1299 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGATGGGGAGAGGGCTGTGCCC 1358

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1359 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 1418
```

```
Qy        737 ATGTCCCCAAAAAGCCGTG 755
              |||||||||||||||||||
Db       1419 ATGTCCCCAAAAAGCCGTG 1437

RESULT 10
US-10-287-226-89
; Sequence 89, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 89
;   LENGTH: 828
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(825)
US-10-287-226-89

  Query Match             90.8%;  Score 700;  DB 17;  Length 828;
  Best Local Similarity   96.6%;  Pred. No. 1.7e-184;
  Matches  715;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | || || ||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         89 GCATTGTTGGGGGGCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 148

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        149 TCCGCGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 208

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              | ||||| |||   ||||||| ||||||| ||||||||||||||||||||||||||||||
Db        209 TAACCGCGGCGCACTGCGTGGAACCGGACATCAAGGATCTGGCCGCCCTCAGGGTGCAAC 268

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        269 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 328

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||| |  ||||| ||| |||||||||||||||||||||||||||||||
Db        329 ACCCACAGTTCTACATCATCCAGACCGGGGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 388

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | |||||  ||||||||||| ||||||||||||| |||||||||||||| ||||||||||
Db        389 CCGTGAACATCTCCAGCCACATCCACACGGTCACGCTGCCCCCTGCCTCGGAGACCTTCC 448

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||| ||||||||| ||||||||||||
Db        449 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGACGTGGACAATAATGAGCGCCTCC 508

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        509 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 568

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        569 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 628

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        629 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 688

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        689 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 748

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        749 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 808

Qy        737 ATGTCCCCAAAAAGCCGTGA 756
              ||||||||||||||||||||
Db        809 ATGTCCCCAAAAAGCCGTGA 828

RESULT 11
US-10-352-684A-45
; Sequence 45, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
;  APPLICANT: Millennium Pharmaceuticals Inc.
```

```
;  APPLICANT:  Carroll, Joseph M.
;  APPLICANT:  Healy, Aileen
;  APPLICANT:  Weich, Nadine S.
;  APPLICANT:  Kelly, Louise M.
;  TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
;  TITLE OF INVENTION:  HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
;  TITLE OF INVENTION:  15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
;  TITLE OF INVENTION:  1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
;  TITLE OF INVENTION:  15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
;  FILE REFERENCE: MPI02-019P1RNOMNIM
;  CURRENT APPLICATION NUMBER: US/10/352,684A
;  CURRENT FILING DATE:  2003-01-28
;  PRIOR APPLICATION NUMBER: US 60/354,333
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: US 60/360,258
;  PRIOR FILING DATE: 2002-02-28
;  PRIOR APPLICATION NUMBER: US 60/364,476
;  PRIOR FILING DATE: 2002-03-15
;  PRIOR APPLICATION NUMBER: US 60/375,626
;  PRIOR FILING DATE: 2002-04-26
;  PRIOR APPLICATION NUMBER: US 60/386,494
;  PRIOR FILING DATE: 2002-06-06
;  PRIOR APPLICATION NUMBER: US 60/390,965
;  PRIOR FILING DATE: 2002-06-24
;  PRIOR APPLICATION NUMBER: US 60/392,480
;  PRIOR FILING DATE: 2002-06-28
;  PRIOR APPLICATION NUMBER: US 60/394,128
;  PRIOR FILING DATE: 2002-07-03
;  PRIOR APPLICATION NUMBER: US 60/399,783
;  PRIOR FILING DATE: 2002-07-31
;  PRIOR APPLICATION NUMBER: US 60/403,221
;  PRIOR FILING DATE: 2002-08-13
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 62
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 45
;   LENGTH: 1154
;   TYPE: DNA
;   ORGANISM: Homo Sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (18)...(845)
US-10-352-684A-45

  Query Match             89.0%;  Score 686.2;  DB 16;  Length 1154;
  Best Local Similarity   94.3%;  Pred. No. 1.2e-180;
  Matches  712;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy          17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         106 GTATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 165

Qy          77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
               ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         166 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 225

Qy         137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
               ||||||| |||   ||| ||||||||||||||||||||||||| |||||||||||||||
Db         226 TGACCGCGGCGCACTGCCTGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAAC 285

Qy         197 TGCGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
               ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db         286 TGCGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGC 345

Qy         257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
               ||||||||||||||| |  |||||  |||||||| |||||||||||||||||||||||||
Db         346 ACCCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGC 405

Qy         317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
               | |||||  ||||||||| ||||||||||||||  |||||||||||||| ||||||||||
Db         406 CCGTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCC 465

Qy         377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
               |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db         466 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCC 525

Qy         437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
               ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         526 CACCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 585

Qy         497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
               |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db         586 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGC 645

Qy         557 TGTGTGCCGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
               |||||||||||||| || ||||||||| ||| |||||||||| ||||||||||||||||
Db         646 TGTGTGCCGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGGCCCCTGGTGT 705

Qy         617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
               |||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||
Db         706 GCAAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCC 765

Qy         677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         766 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 825

Qy         737 ATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
               ||||||||||||||||||||  | | |   |  ||
Db         826 ATGTCCCCAAAAAGCCGTGAGTCAGGCCTGGGTGT 860

RESULT 12
US-10-116-802-240
; Sequence 240, Application US/10116802
; Publication No. US20030065157A1
; GENERAL INFORMATION:
;  APPLICANT: Amy Lasek
;  TITLE OF INVENTION: GENES EXPRESSED IN LUNG CANCER
;  FILE REFERENCE: PA-0045 US
;  CURRENT APPLICATION NUMBER: US/10/116,802
;  CURRENT FILING DATE:  2002-04-04
;  PRIOR APPLICATION NUMBER: 60/281,593
;  PRIOR FILING DATE: 2001-04-04
;  NUMBER OF SEQ ID NOS: 519
;  SOFTWARE: PERL Program
; SEQ ID NO 240
;   LENGTH: 1158
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 378633.40
US-10-116-802-240

  Query Match             89.0%;  Score 686.2;  DB 13;  Length 1158;
  Best Local Similarity   94.9%;  Pred. No. 1.2e-180;
  Matches  709;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy          17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         106 GTATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 165

Qy          77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
               ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         166 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 225

Qy         137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
               ||||||| |||   ||| ||||||||||||||||||||||||| |||||||||||||||
Db         226 TGACCGCGGCGCACTGCCTGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAAC 285

Qy         197 TGCGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
               ||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db         286 TGCGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGC 345

Qy         257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
               ||||||||||||||| |  |||||  |||||||| |||||||||||||||||||||||||
```

```
Db        346 ACCCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGC 405

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | |||||  ||||||||| ||||||||||||||  |||||||||||||| ||||||||||
Db        406 CCGTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCC 465

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        466 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCC 525

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        526 CACCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 585

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        586 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGC 645

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||| || ||||||||| ||| |||||||||| ||||||||||||||||
Db        646 TGTGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGGCCCCTGGTGT 705

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              |||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||
Db        706 GCAAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCC 765

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 825

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        826 ATGTCCCCAAAAAGCCGTGAGTCAGGC 852


RESULT 13
US-10-287-226-87
; Sequence 87, Application US/10287226
; Publication No. US20040086875A1
; GENERAL INFORMATION:
;  APPLICANT: Agee, Michele L.,
;  APPLICANT:  Alsobrook, John P.,
;  APPLICANT:  Berghs, Constance,
;  APPLICANT:  Boldog, Ference,
;  APPLICANT:  Burgess, Catherine E.,
;  APPLICANT:  Chant, John S.,
;  APPLICANT:  Chaudhuri, Amitabha,
;  APPLICANT:  DiPippo, Vincent A.,
;  APPLICANT:  Edinger, Shlomit R.,
;  APPLICANT:  Eisen, Andrew,
;  APPLICANT:  Ellerman, Karen,
;  APPLICANT:  Gangolli, Esha A.,
;  APPLICANT:  Gorman, Linda,
;  APPLICANT:  Gerlach, Valerie,
;  APPLICANT:  Ji, Weizhen,
;  APPLICANT:  Kekuda, Ramesh,
;  APPLICANT:  Khramtsov, Nikolai,
;  APPLICANT:  Li, Li,
;  APPLICANT:  Malyankar, Uriel M.,
;  APPLICANT:  MacDougall, John R.,
;  APPLICANT:  Mezes, Peter S.
;  APPLICANT:  Miller, Charles E.,
;  APPLICANT:  Millet, Isabelle,
;  APPLICANT:  Ooi, Chean Eng,
;  APPLICANT:  Ort, Tatiana,
;  APPLICANT:  Padigaru, Muralidhara,
;  APPLICANT:  Patturajan, Meera,
;  APPLICANT:  Rastelli, Luca,
;  APPLICANT:  Rieger, Daniel K.,
;  APPLICANT:  Rothenberg, Mark E.,
;  APPLICANT:  Shenoy, Suresh G.,
;  APPLICANT:  Spaderna, Steven K.,
;  APPLICANT:  Spytek, Kimberley A.,
```

```
;  APPLICANT:  Taupier, Jr., Raymond J.,
;  APPLICANT:  Vernet, Corine A.M.,
;  APPLICANT:  Zerhusen, Bryan D.,
;  APPLICANT:  Zhong, Mei
;  TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;  FILE REFERENCE: 21402-480C
;  CURRENT APPLICATION NUMBER: US/10/287,226
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: 60/334,421
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,392
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/360,148
;  PRIOR FILING DATE: 2002-02-27
;  PRIOR APPLICATION NUMBER: 60/364,000
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/404,821
;  PRIOR FILING DATE: 2002-08-20
;  PRIOR APPLICATION NUMBER: 60/334,526
;  PRIOR FILING DATE: 2001-11-30
;  PRIOR APPLICATION NUMBER: 60/354,409
;  PRIOR FILING DATE: 2002-02-04
;  PRIOR APPLICATION NUMBER: 60/364,227
;  PRIOR FILING DATE: 2002-03-13
;  PRIOR APPLICATION NUMBER: 60/334,027
;  PRIOR FILING DATE: 2001-11-28
;  PRIOR APPLICATION NUMBER: 60/331,641
;  PRIOR FILING DATE: 2001-11-20
;  Remaining Prior Application data removed - See File Wrapper or PALM.
;  NUMBER OF SEQ ID NOS: 673
;  SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 87
;   LENGTH: 858
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (8)..(844)
US-10-287-226-87

  Query Match             87.9%;  Score 678;  DB 17;  Length 858;
  Best Local Similarity   96.3%;  Pred. No. 2.2e-178;
  Matches  722;  Conservative    0;  Mismatches   10;  Indels   18;  Gaps    2;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        120 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 179

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        180 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 239

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        240 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 299

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        300 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 359

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        360 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 419

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        420 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 479

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              |||||||||||||||||||||||||||||||||||||||||               ||||
Db        480 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTG---------------CTCC 524

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         525 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 584

Qy         497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         585 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 644

Qy         557 TGTGTGCCGGGAACACCCGGAGGGACTCATGC---CAGGGCGACTCCGGAGGGCCCCTGG 613
               ||||||||||||||||||||||||||||||||   |||||||||||||||||||||||||
Db         645 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGCAGGGCGACTCCGGAGGGCCCCTGG 704

Qy         614 TGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTG 673
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         705 TGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTG 764

Qy         674 CCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACC 733
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         765 CCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACC 824

Qy         734 ACTATGTCCCCAAAAAGCCGTGAAGCGGCC 763
               |||||||||||||||||||||||  | | |
Db         825 ACTATGTCCCCAAAAAGCCGTGAGTCAGGC 854

RESULT 14
US-10-275-505-22
; Sequence 22, Application US/10275505
; Publication No. US20040081961A1
; GENERAL INFORMATION:
;  APPLICANT: INCYTE GENOMICS, INC.
;  APPLICANT:  DELEGEANE, Angelo M.; LAL, Preeti G.
;  APPLICANT:  HAFALIA, April J.A.; PATTERSON, Chandra
;  APPLICANT:  WALIA, Narinder K.; KEARNEY, Liam
;  APPLICANT:  TRIBOULEY, Catherine M.; KHAN, Farrah A.
;  APPLICANT:  YAO, Monique G.; BAUGHN, Mariah R.
;  APPLICANT:  AZIMZAI, Yalda; ELLIOTT, Vicki S.
;  APPLICANT:  NGUYEN, Danniel B.; GANDHI, Ameena R.
;  APPLICANT:  YANG, Junming; HERNANDEZ, Roberto
;  APPLICANT:  POLICKY, Jennifer L.; LU, Dyung Aina M.
;  APPLICANT:  REDDY, Roopa M.; YUE, Henry
;  APPLICANT:  TANG, Y. Tom
;  TITLE OF INVENTION: PROTEASES
;  FILE REFERENCE: PI-0085 USN
;  CURRENT APPLICATION NUMBER: US/10/275,505
;  CURRENT FILING DATE:  2002-11-04
;  PRIOR APPLICATION NUMBER: PCT/US01/14651
;  PRIOR FILING DATE: 2001-05-04
;  PRIOR APPLICATION NUMBER: 60/209,402
;  PRIOR FILING DATE: 2000-06-01
;  PRIOR APPLICATION NUMBER: 60/207,477
;  PRIOR FILING DATE: 2000-05-25
;  PRIOR APPLICATION NUMBER: 60/205,803
;  PRIOR FILING DATE: 2000-05-17
;  PRIOR APPLICATION NUMBER: 60/203,566
;  PRIOR FILING DATE: 2000-05-11
;  PRIOR APPLICATION NUMBER: 60/202,082
;  PRIOR FILING DATE: 2000-05-04
;  NUMBER OF SEQ ID NOS: 28
;  SOFTWARE: PERL Program
; SEQ ID NO 22
;   LENGTH: 1187
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte ID No: 3897384CB1
US-10-275-505-22

  Query Match             60.1%;  Score 463;  DB 17;  Length 1187;
  Best Local Similarity   76.4%;  Pred. No. 1.2e-118;
  Matches  568;  Conservative    0;  Mismatches  175;  Indels    0;  Gaps    0;

Qy          14 AAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGA 73
```

```
               || | || || ||||| |||||||| |   ||| |||||||||||||||||||||||||
Db         258 AAGGCATTGTGGGGGGACAGGAGGCACATGGGAACAAGTGGCCCTGGCAGGTGAGCCTGC 317

Qy          74 GAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGG 133
               | | | | | | | ||||||||||| |||||||| ||||||||||||||||| |||||||
Db         318 GTGCCAATGACACCTACTGGATGCATTTCTGCGGTGGCTCCCTCATCCACCCACAGTGGG 377

Qy         134 TGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGC 193
               |||| || || ||    || ||||||||||| ||    || |   |     |||| || |
Db         378 TGCTCACTGCGGCACACTGTGTGGGACCGGATGTTGCTGACCCCAACAAGGTCAGAGTAC 437

Qy         194 AACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCG 253
               | || ||  ||||| ||||||| ||||| ||||| ||| || | || |||  |||||||
Db         438 AGCTCCGTAAGCAGTACCTCTATTACCATGACCACCTGATGACTGTGAGCCAGATCATCA 497

Qy         254 TGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGG 313
                 |||||  | ||||||| || ||||   || || ||||| ||||||||| | ||
Db         498 CACACCCCGACTTCTACATCGTCCAGGATGGGGCAGACATTGCCCTGCTGAAACTCACAA 557

Qy         314 AGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCT 373
               | || |||||  | ||   | | |||||| | ||| |||| || ||||||||||||||||
Db         558 ACCCTGTGAACATTTCTGACTATGTCCACCCTGTCCCCCTACCTCCTGCCTCAGAGACCT 617

Qy         374 TCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCC 433
               ||||| | || | |  ||||||||| || |||||||| |  |  ||||||| ||    ||
Db         618 TCCCCTCAGGAACGTTGTGCTGGGTGACAGGCTGGGGTAACATCGACAATGGTGTAAACC 677

Qy         434 TCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTG 493
               | || || ||||||||| ||||| ||||| | || ||||| || ||||||||| ||||||
Db         678 TGCCGCCACCATTTCCTTTGAAGGAGGTGCAAGTTCCCATTATAGAAAACCACCTTTGTG 737

Qy         494 ACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACA 553
               ||   || || |||   ||   |  ||| || ||| | |||| ||| ||||| || ||||
Db         738 ACTTGAAGTATCACAAAGGTCTCATCACAGGTGACAATGTCCACATTGTCCGAGATGACA 797

Qy         554 TGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGG 613
               |||||||||| |||||     |    ||||| |||||||||||||||||||| || ||||
Db         798 TGCTGTGTGCTGGGAATGAAGGACATGACTCCTGCCAGGGCGACTCCGGAGGACCTCTGG 857

Qy         614 TGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTG 673
               | ||||||||  | | ||||||||||||||| ||||||||||||||||| ||||||||||
Db         858 TCTGCAAGGTAGAAGACACCTGGCTGCAGGCAGGCGTGGTCAGCTGGGGTGAGGGCTGTG 917

Qy         674 CCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACC 733
               | ||||||||| ||||||||||||||||||| |||||||| |||||||||||||||||||
Db         918 CACAGCCCAACAGGCCTGGCATCTACACCCGGGTCACCTATTACTTGGACTGGATCCACC 977

Qy         734 ACTATGTCCCCAAAAAGCCGTGA 756
               |||||||||||||  |    |||
Db         978 ACTATGTCCCCAAGGACTTCTGA 1000

RESULT 15
US-09-813-432-11
; Sequence 11, Application US/09813432
; Publication No. US20030148485A1
; GENERAL INFORMATION:
;  APPLICANT: Taupier Jr., Raymond J
;  APPLICANT:  Majmuder, Kamud
;  APPLICANT:  Spaderna, Steven K
;  APPLICANT:  Smithson, Glenda
;  APPLICANT:  Mezes, Peter S
;  APPLICANT:  Vernet, Corine A. M.
;  TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same
;  FILE REFERENCE: 15966-729
;  CURRENT APPLICATION NUMBER: US/09/813,432
;  CURRENT FILING DATE:  2001-03-20
;  PRIOR APPLICATION NUMBER: 60/190,835
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,768
;  PRIOR FILING DATE: 2000-03-20
;  PRIOR APPLICATION NUMBER: 60/190,972
;  PRIOR FILING DATE: 2000-03-22
```

```
;  PRIOR APPLICATION NUMBER: 60/191,199
;  PRIOR FILING DATE: 2000-03-22
;  PRIOR APPLICATION NUMBER: 60/191,947
;  PRIOR FILING DATE: 2000-03-24
;  PRIOR APPLICATION NUMBER: 60/192,665
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,657
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,984
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,664
;  PRIOR FILING DATE: 2000-03-28
;  PRIOR APPLICATION NUMBER: 60/192,836
;  PRIOR FILING DATE: 2000-03-29
;  PRIOR APPLICATION NUMBER: 60/193,843
;  PRIOR FILING DATE: 2000-03-31
;  NUMBER OF SEQ ID NOS: 78
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
;   LENGTH: 846
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-813-432-11

  Query Match             37.3%;  Score 287.4;  DB 10;  Length 846;
  Best Local Similarity   62.6%;  Pred. No. 6.9e-70;
  Matches  469;  Conservative    0;  Mismatches  271;  Indels    9;  Gaps    1;

Qy          17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
               | || || ||||| ||  | ||||||  | | |||||||| |||||||| |||||||| |
Db          94 GCATTGTGGGGGGCCACAATGCCCCCCCGGGGAAGTGGCCGTGGCAGGTCAGCCTGAGGG 153

Qy          77 TCCACGGCC---------CATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCC 127
               || || ||          | | ||||  |||| ||||  ||||||||||||||||||||||
Db         154 TCTACAGCTACCACTGGGCCTCCTGGGCGCACATCTGTGGGGGCTCCCTCATCCACCCCC 213

Qy         128 AGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCA 187
               ||||||||||||| || ||    ||| |      | | | ||  || | | ||  |  |
Db         214 AGTGGGTGCTGACTGCTGCCCACTGCATTTTCTGGAAGGACACCGACCCGTCCATCTACC 273

Qy         188 GGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGA 247
               || | ||     ||||   | | ||||||           |||||||   |||||| |||
Db         274 GGATCCACGCTGGGGACGTGTATCTCTACGGGGGCCGGGGGCTGCTGAACGTCAGCCGGA 333

Qy         248 TCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGC 307
               ||||||| |||||  | |    ||| ||   | | || |||||  | ||||||||  |||
Db         334 TCATCGTCCACCCCAACTATGTCACTGCGGGGCTGGGTGCGGATGTGGCCCTGCTCCAGC 393

Qy         308 TGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAG 367
               ||| |   ||  |||  |   ||    | |||   |||||||  ||  |||| | |||
Db         394 TGGTGAGCCCCATGATCGGAGCCGCTAATGTCAGGACGGTCAAGCTCTCCCCGGTCTCGC 453

Qy         368 AGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATG 427
                |    || |||||  |   | ||||||||| ||||||||||| |   |    |     |
Db         454 TGGAGCTCACCCCGAAGGACCAGTGCTGGGTGACTGGCTGGGGAGCGATCAGGATGTTCG 513

Qy         428 AGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACA 487
               ||   || || ||||| |  |  ||| |||||| ||  || |   |  |||| |||  |
Db         514 AGTCGCTGCCGCCGCCCTACCGCCTGCAGCAGGCGAGTGTGCAGGTGCTGGAGAACGCCG 573

Qy         488 TTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTG 547
               | |||||       |||| |   | | |          | ||||   |   || |||  |
Db         574 TCTGTGAGCAGCCCTACCGCAACGCCTCAGGGCACACTGGCGACCGGCAGCTCATCCTGG 633

Qy         548 ACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGC 607
               | ||||||||||||||||| | |    |  | ||||| || ||||| |||||||| || |
Db         634 ATGACATGCTGTGTGCCGGCAGCGAGGGCCGAGACTCCTGTCAGGGTGACTCCGGCGGCC 693

Qy         608 CCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGG 667
               | ||||| |||| | ||   ||  ||||||  | || ||| ||||||||||||||| | |
Db         694 CTCTGGTCTGCAGGCTGCGGGGGTCCTGGCGCCTGGTGGGGGTGGTCAGCTGGGGCTACG 753

Qy         668 GCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGA 727
               ||||| ||| ||   ||   || ||| |||||||||  |||     |||| ||  |||||
Db         754 GCTGTACCCTGCGGGACTTTCCCGGCGTCTACACCCACGTCCAGATCTACGTGCTCTGGA 813

Qy         728 TCCACCACTATGTCCCCAAAAAGCCGTGA 756
               |||  ||  | |||    |   ||| |||
Db         814 TCCTGCAGCAAGTCGGGGAGTTGCCCTGA 842

Search completed: July 22, 2004, 15:20:33
Job time : 572 secs
```

```
                 GenCore version 5.1.6
          Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         July 22, 2004, 10:46:50 ; Search time 90 Seconds
                                          (without alignments)
                                          4754.079 Million cell updates/sec

Title:          US-09-598-982-20
Perfect score:  771
Sequence:       1 gggcccctcgagaaaagaat..........cgtgaagcggccgccgtcgt 771

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters:      1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                   SUMMARIES
                 %
Result           Query
  No.   Score    Match  Length DB  ID                      Description
---------------------------------------------------------------------------
    1   764.6    99.2     771  3  US-09-079-970A-4          Sequence 4, Appli
    2     731    94.8    1128  2  US-09-016-366A-20         Sequence 20, Appl
    3     731    94.8    1128  2  US-08-978-404B-15         Sequence 15, Appl
    4     731    94.8    1137  2  US-09-016-366A-18         Sequence 18, Appl
    5     731    94.8    1137  2  US-08-978-404B-13         Sequence 13, Appl
    6   728.6    94.5     735  3  US-09-079-970A-1          Sequence 1, Appli
    7   726.2    94.2    1081  2  US-09-016-366A-22         Sequence 22, Appl
    8   726.2    94.2    1081  2  US-08-978-404B-17         Sequence 17, Appl
    9   686.2    89.0    1154  2  US-09-016-366A-16         Sequence 16, Appl
   10   686.2    89.0    1154  2  US-08-978-404B-11         Sequence 11, Appl
   11   477.2    61.9    1219  2  US-08-978-404B-7          Sequence 7, Appli
   12   469.6    60.9    1108  2  US-09-016-366A-14         Sequence 14, Appl
   13   469.6    60.9    1108  2  US-08-978-404B-20         Sequence 20, Appl
   14     463    60.1    1031  2  US-08-978-404B-1          Sequence 1, Appli
   15   445.6    57.8    1103  2  US-09-016-366A-24         Sequence 24, Appl
   16     444    57.6    1097  2  US-08-978-404B-4          Sequence 4, Appli
   17     371    48.1    2259  2  US-08-845-998-3           Sequence 3, Appli
   18     371    48.1    2259  3  US-09-206-537-3           Sequence 3, Appli
   19     371    48.1    2259  3  US-09-430-854-3           Sequence 3, Appli
   20   369.4    47.9    2218  2  US-08-845-998-5           Sequence 5, Appli
   21   369.4    47.9    2218  3  US-09-206-537-5           Sequence 5, Appli
   22   369.4    47.9    2218  3  US-09-430-854-5           Sequence 5, Appli
   23   271.8    35.3    1095  2  US-08-978-404B-9          Sequence 9, Appli
   24   201.4    26.1     980  4  US-09-023-942A-30         Sequence 30, Appl
   25   201.4    26.1    1110  4  US-09-386-653A-1          Sequence 1, Appli
   26   199.8    25.9    1212  4  US-09-620-312D-431        Sequence 431, App
   27   193.8    25.1    1130  4  US-09-386-653A-8          Sequence 8, Appli
   28   187.8    24.4    1613  4  US-09-387-375-1           Sequence 1, Appli
   29   181.2    23.5    1130  4  US-09-387-375-8           Sequence 8, Appli
   30     153    19.8    3757  2  US-09-016-366A-13         Sequence 13, Appl
   31     153    19.8    3757  2  US-08-978-404B-19         Sequence 19, Appl
   32     152    19.7     933  4  US-09-023-942A-29         Sequence 29, Appl
   33   143.8    18.7    2397  2  US-08-978-404B-2          Sequence 2, Appli
   34   140.6    18.2    1142  4  US-09-386-642-8           Sequence 8, Appli
   35   140.6    18.2    1169  4  US-09-386-642-7           Sequence 7, Appli
   36   132.2    17.1    1154  4  US-09-636-382A-1          Sequence 1, Appli
   37   131.2    17.0    1165  4  US-09-023-942A-28         Sequence 28, Appl
   38   131.2    17.0    1378  4  US-09-907-794A-262        Sequence 262, App
   39   131.2    17.0    1378  4  US-09-905-125A-262        Sequence 262, App
   40   131.2    17.0    1378  4  US-09-902-775A-262        Sequence 262, App
   41   131.2    17.0    1430  4  US-09-386-629-1           Sequence 1, Appli
   42   129.2    16.8    1166  4  US-09-386-629-2           Sequence 2, Appli
   43     124    16.1    1171  4  US-09-636-382A-14         Sequence 14, Appl
   44   122.8    15.9     942  4  US-09-636-382A-3          Sequence 3, Appli
   45     117    15.2    1100  4  US-09-907-794A-256        Sequence 256, App

                                  ALIGNMENTS

RESULT 1
US-09-079-970A-4
; Sequence 4, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  771 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  7..753
;    FEATURE:
;      NAME/KEY:  misc_signal
```

```
;      LOCATION:  7..18
US-09-079-970A-4

  Query Match             99.2%;  Score 764.6;  DB 3;  Length 771;
  Best Local Similarity   99.5%;  Pred. No. 4.1e-166;
  Matches  767;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 2
US-09-016-366A-20
; Sequence 20, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-20

  Query Match             94.8%;  Score 731;  DB 2;  Length 1128;
  Best Local Similarity   98.7%;  Pred. No. 2.1e-158;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 146

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 206

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 266

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 326

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 386

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 446

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 506

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 566

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 626
```

```
Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 686

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 746

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 806

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        807 ATGTCCCCAAAAAGCCGTGAGTCAGGC 833


RESULT 3
US-08-978-404B-15
; Sequence 15, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1128 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-15

  Query Match             94.8%;  Score 731;  DB 2;  Length 1128;
  Best Local Similarity   98.7%;  Pred. No. 2.1e-158;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 146

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 206
```

```
Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 266

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 326

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 386

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 446

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 506

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 566

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 626

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 686

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 746

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 806

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        807 ATGTCCCCAAAAAGCCGTGAGTCAGGC 833


RESULT 4
US-09-016-366A-18
; Sequence 18, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
```

```
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-18

  Query Match             94.8%;  Score 731;  DB 2;  Length 1137;
  Best Local Similarity   98.7%;  Pred. No. 2.1e-158;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 144

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 204

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 264

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 324

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 384

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 444

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 504

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 564

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 624

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 684

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 744

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 804

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        805 ATGTCCCCAAAAAGCCGTGAGTCAGGC 831
```

```
RESULT 5
US-08-978-404B-13
; Sequence 13, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1137 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-13

  Query Match             94.8%;  Score 731;  DB 2;  Length 1137;
  Best Local Similarity   98.7%;  Pred. No. 2.1e-158;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         85 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 144

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        145 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 204

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        205 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 264

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        265 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 324

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        325 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 384

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        385 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 444
```

```
Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        445 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 504

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        505 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 564

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        565 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 624

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        625 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 684

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        685 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 744

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        745 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 804

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        805 ATGTCCCCAAAAAGCCGTGAGTCAGGC 831


RESULT 6
US-09-079-970A-1
; Sequence 1, Application US/09079970A
; Patent No. 6274366
;  GENERAL INFORMATION:
;    APPLICANT:  Maffitt, Mark A.
;    APPLICANT:  Niles, Andrew L.
;    APPLICANT:  Haak-Frendscho, Mary
;    TITLE OF INVENTION:  Enzymatically-Active Recombinant Human
;    TITLE OF INVENTION:  Beta-Tryptase and Method of Making Same
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Intellectual Property Department
;      STREET:  8000 Excelsior Drive, Suite 401
;      CITY:  Madison
;      STATE:  WISCONSIN
;      COUNTRY:  U.S.A.
;      ZIP:  53717-1914
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/079,970A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Leone, Joseph T.
;      REGISTRATION NUMBER:  37,170
;      REFERENCE/DOCKET NUMBER:  34506.073
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (608) 831-2100
;      TELEFAX:  (608) 831-2106
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  735 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..735
US-09-079-970A-1

  Query Match             94.5%;  Score 728.6;  DB 3;  Length 735;
  Best Local Similarity   99.5%;  Pred. No. 6.8e-158;
  Matches  731;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         19 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy         79 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy        139 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 198
              ||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy        199 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy        259 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy        319 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy        379 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy        439 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy        499 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 558
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy        559 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 618
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy        619 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 678
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy        679 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 738
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy        739 GTCCCCAAAAAGCCG 753
              |||||||||||||||
Db        721 GTCCCCAAAAAGCCG 735


RESULT 7
US-09-016-366A-22
; Sequence 22, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
```

```
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  22:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-22

  Query Match             94.2%;  Score 726.2;  DB 2;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 2.6e-157;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812

RESULT 8
US-08-978-404B-17
; Sequence 17, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  17:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1081 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-17

  Query Match             94.2%;  Score 726.2;  DB 2;  Length 1081;
  Best Local Similarity   98.3%;  Pred. No. 2.6e-157;
  Matches  734;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 125
```

```
Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        126 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 185

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        186 TGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 245

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 305

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        306 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 365

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        366 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 425

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 485

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        486 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 545

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        546 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 605

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        606 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 665

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        666 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 725

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        726 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 785

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        786 ATGTCCCCAAAAAGCCGTGAGTCAGGC 812

RESULT 9
US-09-016-366A-16
; Sequence 16, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
```

```
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1154 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-16

  Query Match             89.0%;  Score 686.2;  DB 2;  Length 1154;
  Best Local Similarity   94.3%;  Pred. No. 3.7e-148;
  Matches  712;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 GTATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 165

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 225

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||| |||||||||||||||||||||||||| |||||||||||||||
Db        226 TGACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAAC 285

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        286 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGC 345

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||| |  |||||  |||||||| |||||||||||||||||||||||||
Db        346 ACCCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGC 405

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | |||||  ||||||||| ||||||||||||||  |||||||||||||| ||||||||||
Db        406 CCGTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCC 465

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db        466 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCC 525

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        526 CACCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 585

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db        586 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGC 645

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||| || ||||||||| ||| |||||||||| ||||||||||||||||
Db        646 TGTGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGGCCCCTGGTGT 705

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              |||||||||||||||||||||| |||||||||||||||||||||| ||||||||||||||
Db        706 GCAAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCC 765

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 825
```

```
Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              ||||||||||||||||||||  | | |   |  ||
Db        826 ATGTCCCCAAAAAGCCGTGAGTCAGGCCTGGGTGT 860


RESULT 10
US-08-978-404B-11
; Sequence 11, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1154 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-11

  Query Match             89.0%;  Score 686.2;  DB 2;  Length 1154;
  Best Local Similarity   94.3%;  Pred. No. 3.7e-148;
  Matches  712;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        106 GTATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 165

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db        166 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 225

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||| |||||||||||||||||||||||||| |||||||||||||||
Db        226 TGACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAAC 285

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              |||||||||||||||||||||||||||||||||||||||| |||||||||||||||||||
Db        286 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCAGTCAGCAGGATCATCGTGC 345

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||| |  |||||  |||||||| |||||||||||||||||||||||||
Db        346 ACCCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGC 405

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | |||||  ||||||||| ||||||||||||||  |||||||||||||| ||||||||||
Db        406 CCGTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCC 465

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        466 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCC 525

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        526 CACCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 585

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db        586 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGC 645

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||| || ||||||||| ||| |||||||||| ||||||||||||||||
Db        646 TGTGTGCCGGGAACAGCCAGAGGGACTCCTGCAAGGGCGACTCTGGAGGGCCCCTGGTGT 705

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              |||||||||||||||||||||| ||||||||||||||||||||||| ||||||||||||||
Db        706 GCAAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGACGAGGGCTGTGCCC 765

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        766 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 825

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              ||||||||||||||||||||  | | |   |  ||
Db        826 ATGTCCCCAAAAAGCCGTGAGTCAGGCCTGGGTGT 860


RESULT 11
US-08-978-404B-7
; Sequence 7, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1219 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-7

  Query Match             61.9%;  Score 477.2;  DB 2;  Length 1219;
  Best Local Similarity   77.1%;  Pred. No. 2.2e-100;
  Matches  581;  Conservative    0;  Mismatches  173;  Indels    0;  Gaps    0;

Qy          6 CCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGT 65
              ||  |||    | || || ||||| |||||||| ||  ||| ||||||||||||||||||
Db        260 CCAAGAGTGGGGCATTGTTGGGGGACAGGAGGCACCTGGGAACAAGTGGCCCTGGCAGGT 319

Qy         66 GAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCC 125
              ||||||  | | | | |   | ||||||| ||| |||||||| |||||||||||||||||
Db        320 GAGCCTTCGTGCCAATGAAACCTACTGGAGGCATTTCTGCGGCGGCTCCCTCATCCACCC 379

Qy        126 CCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCT 185
               ||||||||||| ||||| ||    || |||||||||    |    ||||    |    |
Db        380 ACAGTGGGTGCTCACCGCGGCACACTGTGTGGGACCGACTATTGCTGATCCCAACAAGGT 439

Qy        186 CAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAG 245
              ||| || || || ||  ||||| ||||||| ||||| ||||| |||||| | || ||| |
Db        440 CAGAGTACAGCTTCGAAAGCAGTACCTCTATTACCACGACCACCTGCTGGCTGTGAGCCG 499

Qy        246 GATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGA 305
              |||||||   |||||    |||||  || ||||||  || |||||||||||||| || ||
Db        500 GATCATCACACACCCGACATTCTATGCCACCCAGAATGGGGCGGACATCGCCCTACTTGA 559

Qy        306 GCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTC 365
              |||  || | || || ||  | |||||||| |||||| | ||| ||||||| ||||||||
Db        560 GCTCAAGAACCCTGTAAACATTTCCAGCCATGTCCACCCCGTCTCCCTGCCTCCTGCCTC 619

Qy        366 AGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGA 425
              ||||||||||||| | || |    ||||||||| || |||||||||  |  | ||||||||
Db        620 AGAGACCTTCCCCTCAGGAACATTGTGCTGGGTGACAGGCTGGGGAAACATCGACAATGA 679

Qy        426 TGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCA 485
              || |  ||| ||||||||||||||  ||||| ||||| | || ||| |  ||||||||||
Db        680 TGTGAGCCTGCCACCGCCATTTCCCTTGAAGGAGGTGCAAGTTCCCGTCGTGGAAAACCA 739

Qy        486 CATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCG 545
                ||||||||   || || |||   || | |||||| || ||| || ||| ||| |||||
Db        740 GCTTTGTGACCTGAAGTATCACAAAGGTGTCTACACAGGGGACAACATCCACATTGTCCG 799

Qy        546 TGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGG 605
               ||||||||||||||||| ||||||    |    ||||| |||||||| |||||||||||
Db        800 AGACGACATGCTGTGTGCTGGGAACGAAGGACACGACTCCTGCCAGGGTGACTCCGGAGG 859

Qy        606 GCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGA 665
               || ||||| |||||||| || || |||||||||||||| || |||||||||||||| ||
Db        860 ACCTCTGGTCTGCAAGGTAAACGGTACCTGGCTGCAGGCAGGTGTGGTCAGCTGGGGTGA 919

Qy        666 GGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTG 725
              ||||||||| | ||||||| |||||||||||||||| || |||||||| |||||||||||
Db        920 GGGCTGTGCTCTGCCCAACAGGCCTGGCATCTACACTCGGGTCACCTATTACTTGGACTG 979

Qy        726 GATCCACCACTATGTCCCCAAAAAGCCGTGAAGC 759
              |||||||| ||||||||||||  |    |||| |
Db        980 GATCCACCGCTATGTCCCCAAGGACTTCTGAATC 1013

RESULT 12
US-09-016-366A-14
; Sequence 14, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    APPLICANT:  Huang, Chifu
;    TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;    TITLE OF INVENTION:  INHIBITORS
;    NUMBER OF SEQUENCES:  65
;    CORRESPONDENCE ADDRESS:
```

```
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/016,366A
;      FILING DATE:  January 30, 1998
;      CLASSIFICATION:  530
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/037,090
;      FILING DATE:  05-FEB-1997
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7093
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-09-016-366A-14

  Query Match             60.9%;  Score 469.6;  DB 2;  Length 1108;
  Best Local Similarity   77.2%;  Pred. No. 1.2e-98;
  Matches  571;  Conservative    0;  Mismatches  169;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| || || || |||||  |   ||| |||||||||||||||||||||||||||
Db        125 GCATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGAT 184

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              |  |       |||||||| || |||||||| ||||| ||||||||||| ||||||||||
Db        185 TTAAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGC 244

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              | || || ||    || ||||||||| || ||||    |       | || ||||||| |
Db        245 TCACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGC 304

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              | || |||||| | || |||||   ||||||||| ||  |  | | | ||||| | ||||
Db        305 TTCGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGC 364

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              |||| || |  ||||| ||| ||   || || ||| | |||||||||||||| ||||  |
Db        365 ACCCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCC 424

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | ||||| ||||||| |||  ||||| |  |  |||||||||||||||| ||||||||||
Db        425 CTGTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCC 484

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              |||| ||||   |||||||||| || |||||||||||  | || ||||| ||||  ||||
Db        485 CCCCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCC 544

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              |||| || | |||||||||||| |||||||| |||||  ||||||||  | | ||||||
Db        545 CACCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACC 604

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                || ||||||  ||||  |||||||||||| ||  | | ||| |||| ||| | |||||
```

```
Db          605 GGAAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGC 664

Qy          557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
                ||||||| || || ||| |||| ||||| ||||||||||| || || ||||| ||||| |
Db          665 TGTGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCT 724

Qy          617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
                |||| ||||| || |||||||||||||| || |||||||||||||| |||||||| || |
Db          725 GCAAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGGTGAGGGCTGCGCAC 784

Qy          677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
                ||||||||  |||||||||||||||||| || || |||||||| ||||||||||||| ||
Db          785 AGCCCAACAAGCCTGGCATCTACACCCGGGTGACATACTACTTAGACTGGATCCACCGCT 844

Qy          737 ATGTCCCCAAAAAGCCGTGA 756
                |||||||  |  |  | |||
Db          845 ATGTCCCTGAGCATTCCTGA 864

RESULT 13
US-08-978-404B-20
; Sequence 20, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  60/032,354
;      FILING DATE:  04-DEC-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plumer, Elizabeth R.
;      REGISTRATION NUMBER:  36,637
;      REFERENCE/DOCKET NUMBER:  B0801/7090
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-720-3500
;      TELEFAX:  617-720-2441
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  20:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1108 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-978-404B-20

  Query Match             60.9%;  Score 469.6;  DB 2;  Length 1108;
  Best Local Similarity   77.2%;  Pred. No. 1.2e-98;
  Matches  571;  Conservative    0;  Mismatches  169;  Indels    0;  Gaps    0;

Qy           17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
                | ||||| || || || |||||  |   ||| ||||||||||||||||||||||||||| 
Db          125 GCATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGAT 184

Qy           77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
                |  |       |||||||| || |||||||| ||||| ||||||||||| ||||||||||
Db          185 TTAAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGC 244

Qy          137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
                | || || ||    || |||||||||| || ||||    |       | || ||||||| |
Db          245 TCACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGC 304

Qy          197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
                | || |||||| | || |||||   |||||||||| ||  |  | | | ||||| | ||||
Db          305 TTCGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGC 364

Qy          257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
                |||| || |  ||||| ||| ||   || || ||| | |||||||||||||| ||||  |
Db          365 ACCCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCC 424

Qy          317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
                | ||||| ||||||| |||  ||||| |  |  ||||||||||||||||| ||||||||||
Db          425 CTGTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCC 484

Qy          377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
                |||| ||||   |||||||||| || |||||||||||  | || ||||| ||||  ||||
Db          485 CCCCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCC 544

Qy          437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
                |||| || | |||||||||||| |||||||| |||||  ||||||||  | | |||||| 
Db          545 CACCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACC 604

Qy          497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                  || ||||||  ||||  ||||||||||| ||  | | ||| |||| ||| | |||||
Db          605 GGAAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGC 664

Qy          557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
                ||||||| || || ||| |||| ||||| ||||||||||| || || ||||| ||||| |
Db          665 TGTGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCT 724

Qy          617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
                |||| ||||| || |||||||||||||| || |||||||||||||| |||||||| || |
Db          725 GCAAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGGTGAGGGCTGCGCAC 784

Qy          677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
                ||||||||  |||||||||||||||||| || || |||||||| ||||||||||||| ||
Db          785 AGCCCAACAAGCCTGGCATCTACACCCGGGTGACATACTACTTAGACTGGATCCACCGCT 844

Qy          737 ATGTCCCCAAAAAGCCGTGA 756
                |||||||  |  |  | |||
Db          845 ATGTCCCTGAGCATTCCTGA 864

RESULT 14
US-08-978-404B-1
; Sequence 1, Application US/08978404B
; Patent No. 5968782
;  GENERAL INFORMATION:
;    APPLICANT:  Stevens, Richard L.
;    TITLE OF INVENTION:  MAST CELL PROTEASE THAT CLEAVES
;    TITLE OF INVENTION:  FIBRINOGEN
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;      STREET:  600 Atlantic Avenue
;      CITY:  Boston
;      STATE:  MA
;      COUNTRY:  U.S.A.
;      ZIP:  02210-2211
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/978,404B
;      FILING DATE:  25-NOV-97
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  60/032,354
;       FILING DATE:  04-DEC-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Plumer, Elizabeth R.
;       REGISTRATION NUMBER:  36,637
;       REFERENCE/DOCKET NUMBER:  B0801/7090
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-720-3500
;       TELEFAX:  617-720-2441
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1031 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-08-978-404B-1

  Query Match             60.1%;  Score 463;  DB 2;  Length 1031;
  Best Local Similarity   76.4%;  Pred. No. 3.8e-97;
  Matches  568;  Conservative    0;  Mismatches  175;  Indels    0;  Gaps    0;

Qy           14 AAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGA 73
                || | || || ||||| |||||||| |   ||| |||||||||||||||||||||||||
Db          106 AAGGCATTGTGGGGGGACAGGAGGCACATGGGAACAAGTGGCCCTGGCAGGTGAGCCTGC 165

Qy           74 GAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGG 133
                | | | | | | | ||||||||||| |||||||| ||||||||||||||||| |||||||
Db          166 GTGCCAATGACACCTACTGGATGCATTTCTGCGGTGGCTCCCTCATCCACCCACAGTGGG 225

Qy          134 TGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGC 193
                |||| || || ||    || ||||||||||| ||    || |    |    |||| || |
Db          226 TGCTCACTGCGGCACACTGTGTGGGACCGGATGTTGCTGACCCCAACAAGGTCAGAGTAC 285

Qy          194 AACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCG 253
                | || ||  ||||| ||||||| ||||| ||||| ||| || | || |||  |||||||
Db          286 AGCTCCGTAAGCAGTACCTCTATTACCATGACCACCTGATGACTGTGAGCCAGATCATCA 345

Qy          254 TGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGG 313
                  |||||  | ||||||| || ||||   || || ||||| ||||||||| | ||
Db          346 CACACCCCGACTTCTACATCGTCCAGGATGGGGCAGACATTGCCCTGCTGAAACTCACAA 405

Qy          314 AGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCT 373
                | || |||||  | ||   | | |||||| | ||| |||| || ||||||||||||||||
Db          406 ACCCTGTGAACATTTCTGACTATGTCCACCCTGTCCCCCTACCTCCTGCCTCAGAGACCT 465

Qy          374 TCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCC 433
                ||||| | || | |  ||||||||| || ||||||||  |  | ||||||| ||    ||
Db          466 TCCCCTCAGGAACGTTGTGCTGGGTGACAGGCTGGGGTAACATCGACAATGGTGTAAACC 525

Qy          434 TCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTG 493
                | || |  ||||||||| ||||| ||||| | || ||||| || ||||||||| ||||||
Db          526 TGCCGCCACCATTTCCTTTGAAGGAGGTGCAAGTTCCCATTATAGAAAACCACCTTTGTG 585

Qy          494 ACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACA 553
                ||   || || |||   ||   |  ||| || ||| | |||| ||| ||||| || ||||
Db          586 ACTTGAAGTATCACAAAGGTCTCATCACAGGTGACAATGTCCACATTGTCCGAGATGACA 645

Qy          554 TGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGG 613
                |||||||||| |||||     |    ||||| |||||||||||||||||||| || ||||
Db          646 TGCTGTGTGCTGGGAATGAAGGACATGACTCCTGCCAGGGCGACTCCGGAGGACCTCTGG 705

Qy          614 TGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTG 673
                | ||||||||  | | ||||||||||||||| ||||||||||||||||| ||||||||||
Db          706 TCTGCAAGGTAGAAGACACCTGGCTGCAGGCAGGCGTGGTCAGCTGGGGTGAGGGCTGTG 765

Qy          674 CCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACC 733
                | ||||||||| ||||||||||||||||||| |||||||| |||||||||||||||||||
Db          766 CACAGCCCAACAGGCCTGGCATCTACACCCGGGTCACCTATTACTTGGACTGGATCCACC 825

Qy          734 ACTATGTCCCCAAAAAGCCGTGA 756
                |||||||||||||  |    |||
Db          826 ACTATGTCCCCAAGGACTTCTGA 848
```

```
RESULT 15
US-09-016-366A-24
; Sequence 24, Application US/09016366A
; Patent No. 5955431
;  GENERAL INFORMATION:
;     APPLICANT:  Stevens, Richard L.
;     APPLICANT:  Huang, Chifu
;     TITLE OF INVENTION:  MAST CELL PROTEASE PEPTIDE
;     TITLE OF INVENTION:  INHIBITORS
;     NUMBER OF SEQUENCES:  65
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Wolf, Greenfield & Sacks, P.C.
;       STREET:  600 Atlantic Avenue
;       CITY:  Boston
;       STATE:  MA
;       COUNTRY:  U.S.A.
;       ZIP:  02210-2211
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/016,366A
;       FILING DATE:  January 30, 1998
;       CLASSIFICATION:  530
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/037,090
;       FILING DATE:  05-FEB-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Plumer, Elizabeth R.
;       REGISTRATION NUMBER:  36,637
;       REFERENCE/DOCKET NUMBER:  B0801/7093
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-720-3500
;       TELEFAX:  617-720-2441
;       TELEX:
;  INFORMATION FOR SEQ ID NO:  24:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1103 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-09-016-366A-24

  Query Match             57.8%;  Score 445.6;  DB 2;  Length 1103;
  Best Local Similarity   75.1%;  Pred. No. 3.7e-93;
  Matches  556;  Conservative    0;  Mismatches  184;  Indels    0;  Gaps    0;

Qy           17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
                | || || || || |  |||||  |    || |||||||||||||||||||||||||||
Db          102 GCATTGTGGGAGGACGAGAGGCTTCTGAAAGTAAGTGGCCCTGGCAGGTGAGCCTGAGAT 161

Qy           77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
                |  |   |   | ||||||||| ||||| || ||||||||||| ||||| ||||||||||
Db          162 TTAAATTCAGCTTCTGGATGCATTTCTGTGGCGGCTCCCTCATTCACCCACAGTGGGTGC 221

Qy          137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
                | || || ||    || ||||||| | || ||||    |  |    | || | || || |
Db          222 TCACTGCGGCACACTGTGTGGGACTGCACATCAAAAGCCCAGAGCTCTTCCGTGTACAGC 281

Qy          197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
                | || ||||||| | || |||||   ||||||||| ||| | || | | ||| | | ||||
Db          282 TTCGTGAGCAGTATCTATACTATGCGGACCAGCTACTGACTGTGAACCGGACCGTTGTGC 341

Qy          257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
                |||| || | |||||| | | ||   || || ||||| |||||||||||||| ||| | |
Db          342 ACCCCCACTACTACACAGTCGAGGATGGGGCAGACATTGCCCTGCTGGAGCTTGAGAACC 401

Qy          317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
```

```
              | ||||| ||||||| |||  ||||| |     |||||||||||||||| ||||||||||
Db        402 CTGTGAATGTCTCCACCCATATCCACCCCACATCCCTGCCCCCTGCCTCGGAGACCTTCC 461

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              || ||||||   | |||||||| || |||||||||||  | || | ||| ||||  ||||
Db        462 CCTCGGGGACTTCTTGCTGGGTAACAGGCTGGGGCGACATTGATAGTGACGAGCCTCTCC 521

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
                || || | |||||||||||| ||||||||||||||  ||||||||  | | |||||
Db        522 TGCCACCTTATCCTCTGAAGCAAGTGAAGGTCCCCATTGTGGAAAACAGCCTGTGTGATC 581

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                || ||||||  ||||  |||||| ||||| || || | ||| ||||  || | |||||
Db        582 GGAAGTACCACACTGGCCTCTACACAGGAGATGATGTTCCCATTGTCCAGGATGGCATGC 641

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||| || || ||| |||| ||||| |||||||| ||||| || || || ||||| |
Db        642 TGTGTGCTGGAAATACCAGGAGCGACTCCTGCCAGGGAGACTCAGGGGGCCCACTGGTCT 701

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              |||| |||||  || ||||||||||| || || ||||||||||||||||||||||| ||
Db        702 GCAAAGTGAAGGGTACCTGGCTGCAAGCAGGAGTGGTCAGCTGGGGCGAGGGCTGCGCAG 761

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              || |||| || |||||||| |||||||| || || |||||| |||||||||| |||| ||
Db        762 AGGCCAATCGTCCTGGCATTTACACCCGGGTGACGTACTACCTGGACTGGATTCACCGCT 821

Qy        737 ATGTCCCCAAAAAGCCGTGA 756
              |||||||  |     | |||
Db        822 ATGTCCCTCAGCGTTCCTGA 841
```

Search completed: July 22, 2004, 13:26:09
Job time : 93 secs

```
                 GenCore version 5.1.6
         Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         July 22, 2004, 09:37:10 ; Search time 522 Seconds
                                           (without alignments)
                                           6274.637 Million cell updates/sec

Title:          US-09-598-982-20
Perfect score:  771
Sequence:       1 gggcccctcgagaaaagaat..........cgtgaagcggccgccgtcgt 771

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters:      6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       N_Geneseq_29Jan04:*
                1:  geneseqn1980s:*
                2:  geneseqn1990s:*
                3:  geneseqn2000s:*
                4:  geneseqn2001as:*
                5:  geneseqn2001bs:*
                6:  geneseqn2002s:*
                7:  geneseqn2003as:*
                8:  geneseqn2003bs:*
                9:  geneseqn2003cs:*
                10:  geneseqn2004s:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                    %
Result              Query
  No.    Score   Match Length DB  ID                        Description
-------------------------------------------------------------------------------
    1      771   100.0     771  6  AAS20775                  Aas20775 DNA encod
    2    769.4    99.8     771  6  AAS20783                  Aas20783 DNA encod
    3    764.6    99.2     771  3  AAZ40175                  Aaz40175 Human bet
    4    764.6    99.2     771  6  AAS20765                  Aas20765 DNA encod
    5      763    99.0     771  6  AAS20763                  Aas20763 DNA encod
    6    761.4    98.8     771  6  AAS20776                  Aas20776 DNA encod
    7    759.8    98.5     771  6  AAS20784                  Aas20784 DNA encod
    8      755    97.9     771  6  AAS20778                  Aas20778 DNA encod
    9      755    97.9     771  6  AAS20777                  Aas20777 DNA encod
   10    753.4    97.7     771  6  AAS20786                  Aas20786 DNA encod
   11    753.4    97.7     771  6  AAS20785                  Aas20785 DNA encod
   12      735    95.3     735  6  AAS20779                  Aas20779 DNA encod
   13    733.4    95.1     735  6  AAS20787                  Aas20787 DNA encod
   14      731    94.8    1128  2  AAV44330                  Aav44330 Human mas
   15      731    94.8    1128  2  AAV42712                  Aav42712 Human mas
   16      731    94.8    1137  2  AAV44329                  Aav44329 Human mas
   17      731    94.8    1137  2  AAV42711                  Aav42711 Human mas
   18      731    94.8    1137  3  AAA34957                  Aaa34957 Human ade
   19      731    94.8    1137  3  AAF21079                  Aaf21079 Human low
   20      731    94.8    1137  7  ABZ96773                  Abz96773 Human nuc
   21      731    94.8    1143  3  AAA34955                  Aaa34955 Human ade
   22      731    94.8    1143  3  AAF21077                  Aaf21077 Human low
   23      731    94.8    1143  7  ABZ96771                  Abz96771 Human nuc
   24      731    94.8    1145  3  AAA34956                  Aaa34956 Human ade
   25      731    94.8    1145  3  AAF21078                  Aaf21078 Human low
   26      731    94.8    1145  7  ABZ96772                  Abz96772 Human nuc
   27      731    94.8   17133  3  AAA34962                  Aaa34962 Human ade
   28      731    94.8   17133  3  AAF21084                  Aaf21084 Human low
   29      731    94.8   17133  7  ABZ96778                  Abz96778 Human nuc
   30    728.6    94.5     735  3  AAZ40172                  Aaz40172 Human bet
   31    728.6    94.5     735  6  AAS20766                  Aas20766 DNA encod
   32      727    94.3     735  6  AAS20760                  Aas20760 DNA encod
   33    726.2    94.2    1081  2  AAV44331                  Aav44331 Human mas
   34    726.2    94.2    1081  2  AAV42713                  Aav42713 Human mas
   35    726.2    94.2    1081  3  AAA34960                  Aaa34960 Human ade
   36    726.2    94.2    1081  3  AAF21082                  Aaf21082 Human low
   37    726.2    94.2    1081  6  ABL62377                  Abl62377 Colon ade
   38    726.2    94.2    1081  6  ABL66816                  Abl66816 Lung canc
   39    726.2    94.2    1081  6  ABL61818                  Abl61818 Colon ade
   40    726.2    94.2    1081  6  ABK35580                  Abk35580 Gene TPS2
   41    726.2    94.2    1081  6  ABK64532                  Abk64532 Human ben
   42    726.2    94.2    1081  7  ABZ96776                  Abz96776 Human nuc
   43    725.4    94.1     735  6  AAS20780                  Aas20780 DNA encod
   44    723.8    93.9     735  6  AAS20788                  Aas20788 DNA encod
   45      719    93.3     735  6  AAS20782                  Aas20782 DNA encod

                                 ALIGNMENTS

RESULT 1
AAS20775
ID   AAS20775 standard; DNA; 771 BP.
XX
AC   AAS20775;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant H44A #1.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12009.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 83-84; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
```

```
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 147 A; 256 C; 244 G; 124 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 771;  DB 6;  Length 771;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-147;
  Matches  771;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 2
AAS20783
ID   AAS20783 standard; DNA; 771 BP.
XX
AC   AAS20783;
```

```
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant H44A #3.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12017.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 104-105; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 147 A; 257 C; 243 G; 124 T; 0 U; 0 Other;

  Query Match             99.8%;  Score 769.4;  DB 6;  Length 771;
  Best Local Similarity   99.9%;  Pred. No. 2.8e-147;
  Matches  770;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy          301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
                ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db          301 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy          661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy          721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 3
AAZ40175
ID   AAZ40175 standard; DNA; 771 BP.
XX
AC   AAZ40175;
XX
DT   18-FEB-2000  (first entry)
XX
DE   Human beta-tryptase coding sequence.
XX
KW   Beta-tryptase; human; DNA expression construct; protein production;
KW   combinatorial library screening; X ray crystallography; antigen;
KW   antibody generation; ss.
XX
OS   Homo sapiens.
XX
PN   WO9960139-A1.
XX
PD   25-NOV-1999.
XX
PF   29-OCT-1998;   98WO-US022994.
XX
PR   15-MAY-1998;   98US-00079970.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffitt MA,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2000-053300/04.
DR   P-PSDB; AAY55011.
XX
PT   New DNA expression construct for production of enzymatically active
PT   recombinant human beta-tryptase.
XX
PS   Disclosure; Page 43-44; 50pp; English.
XX
CC   This sequence encodes the human beta-tryptase. The invention relates to a
CC   DNA expression construct comprising (5' to 3') a promoter linked to a
CC   signal sequence which is linked to a sequence encoding human beta-
CC   tryptase. The DNA construct is useful for transforming host cells to
CC   express, post translationally process and secrete enzymatically active
CC   human tryptase. The method is useful for the production of large amounts
CC   of tryptase with defined specifications. The transformant is useful for
CC   pharmacological studies, combinatorial library screens and X ray
CC   crystallographic studies. The tryptase produced allows for the
CC   development of tryptase agonists and/or antagonists, is useful as an
CC   antigen to generate antihuman tryptase antibodies in various animals, can
CC   be used in screening for compounds which act as tryptase inhibitors,
CC   antagonists, agonists etc. and to assay for the presence of tryptase in
CC   biological or other solutions. Tryptase inhibitors, antagonists, agonists
CC   etc. may be useful as therapeutics. The tryptase does not require any
CC   post-expression or post-purification modifications or manipulations to
CC   initiate tryptase activity and it has enzymatic activity which compares
CC   favourably with cadaveric tryptase. The availability of enzymatically
CC   active tryptase facilitates the large scale screening of combinatorial
CC   libraries for specific tryptase inhibitors as potential therapeutics and
CC   advances the understanding of the biological significance of tryptase in
CC   mast cell mediated diseases. The tryptase can be used to detect low
CC   levels of tryptase
XX
SQ   Sequence 771 BP; 149 A; 256 C; 242 G; 124 T; 0 U; 0 Other;

  Query Match             99.2%;  Score 764.6;  DB 3;  Length 771;
  Best Local Similarity   99.5%;  Pred. No. 2.6e-146;
  Matches  767;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy            1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy           61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy          121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
                ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db          121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy          181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy          301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
```

```
Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 4
AAS20765
ID   AAS20765 standard; DNA; 771 BP.
XX
AC   AAS20765;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding recombinant human beta-II tryptase.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; gene; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   CDS             7..756
FT                   /*tag=  a
FT                   /partial
FT                   /product= "Recombinant beta-II tryptase"
FT                   /note= "This sequence lacks a start codon"
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12007.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 43; Page 75-77; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   The present sequence encodes for recombinant human beta-II tryptase
XX
SQ   Sequence 771 BP; 149 A; 256 C; 242 G; 124 T; 0 U; 0 Other;

  Query Match             99.2%;  Score 764.6;  DB 6;  Length 771;
  Best Local Similarity   99.5%;  Pred. No. 2.6e-146;
  Matches  767;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 5
AAS20763
ID   AAS20763 standard; DNA; 771 BP.
XX
AC   AAS20763;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-I tryptase.
XX
KW   Human; proteolytic tryptase; protease; beta-I tryptase; enzyme; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             7..756
FT                   /*tag=  a
```

```
FT                   /partial
FT                   /product= "Beta-I tryptase"
FT                   /note= "This sequence lacks a start codon"
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12006.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Disclosure; Page 72-74; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   The present sequence encodes for human beta-I tryptase
XX
SQ   Sequence 771 BP; 149 A; 257 C; 241 G; 124 T; 0 U; 0 Other;

  Query Match             99.0%;  Score 763;  DB 6;  Length 771;
  Best Local Similarity   99.4%;  Pred. No. 5.6e-146;
  Matches  766;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 6
AAS20776
ID   AAS20776 standard; DNA; 771 BP.
XX
AC   AAS20776;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant D91A #1.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12010.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 86-87; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
```

```
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 149 A; 256 C; 242 G; 124 T; 0 U; 0 Other;

  Query Match             98.8%;  Score 761.4;  DB 6;  Length 771;
  Best Local Similarity   99.2%;  Pred. No. 1.2e-145;
  Matches  765;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 7
AAS20784
ID   AAS20784 standard; DNA; 771 BP.
XX
AC   AAS20784;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant D91A #3.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12018.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 107-109; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 149 A; 257 C; 241 G; 124 T; 0 U; 0 Other;

  Query Match             98.5%;  Score 759.8;  DB 6;  Length 771;
  Best Local Similarity   99.1%;  Pred. No. 2.5e-145;
  Matches  764;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
```

```
Qy          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
                |||||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGCAATCGCCCTG 300

Qy          301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
                ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db          301 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600

Qy          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy          661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy          721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
                |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
```

```
RESULT 8
AAS20778
ID   AAS20778 standard; DNA; 771 BP.
XX
AC   AAS20778;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant S194A #2.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12012.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 91-92; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 152 A; 255 C; 241 G; 123 T; 0 U; 0 Other;

  Query Match             97.9%;  Score 755;  DB 6;  Length 771;
  Best Local Similarity   98.7%;  Pred. No. 2.4e-144;
  Matches  761;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy            1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy           61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy          121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
                ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db          121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy          181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy          301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
                |||||||||||||||||||||||||||||||||||||||||||||||||| || ||| ||
Db          541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCC 600

Qy          601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
                || || || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 GGCGGACCACTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy          661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
```

```
Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771

RESULT 9
AAS20777
ID   AAS20777 standard; DNA; 771 BP.
XX
AC   AAS20777;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant S194A #1.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12011.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 88-90; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 150 A; 255 C; 241 G; 125 T; 0 U; 0 Other;

  Query Match             97.9%;  Score 755;  DB 6;  Length 771;
  Best Local Similarity   98.7%;  Pred. No. 2.4e-144;
  Matches  761;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
```

```
Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||| || |||||| ||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              || || || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCGGACCTCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771

RESULT 10
AAS20786
ID   AAS20786 standard; DNA; 771 BP.
XX
AC   AAS20786;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant S194A #6.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
```

```
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12020.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 113-114; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 152 A; 256 C; 240 G; 123 T; 0 U; 0 Other;

  Query Match             97.7%;  Score 753.4;  DB 6;  Length 771;
  Best Local Similarity   98.6%;  Pred. No. 5e-144;
  Matches  760;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              |||||||||||||||||||||||||||||||||||||||||||||||||| || ||| ||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAAGGAGACGCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              || || || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCGGACCACTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771


RESULT 11
AAS20785
ID   AAS20785 standard; DNA; 771 BP.
XX
AC   AAS20785;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant S194A #5.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12019.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Claim 7; Page 110-111; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 771 BP; 150 A; 256 C; 240 G; 125 T; 0 U; 0 Other;

  Query Match             97.7%;  Score 753.4;  DB 6;  Length 771;
  Best Local Similarity   98.6%;  Pred. No. 5e-144;
```

```
  Matches  760;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 GGGCCCCTCGAGAAAAGAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGG 60

Qy         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CAGGTGAGCCTGAGAGTCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATC 120

Qy        121 CACCCCCAGTGGGTGCTGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCC 180
              ||||||||||||||||||||||| |||   ||||||||||||||||||||||||||||||
Db        121 CACCCCCAGTGGGTGCTGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCC 180

Qy        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTC 240

Qy        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300

Qy        301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
              ||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db        301 CTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360

Qy        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420

Qy        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480

Qy        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540

Qy        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
              ||||||||||||||||||||||||||||||||||||||||||||||| || |||||| ||
Db        541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGTCAAGGCGACGCC 600

Qy        601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
              || || || |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        601 GGCGGACCTCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660

Qy        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720

Qy        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db        721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCCGCCGTCGT 771

RESULT 12
AAS20779
ID   AAS20779 standard; DNA; 735 BP.
XX
AC   AAS20779;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant H44A #2.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
```

```
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12013.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Example 1c; Page 94-95; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 735 BP; 139 A; 245 C; 231 G; 120 T; 0 U; 0 Other;

  Query Match             95.3%;  Score 735;  DB 6;  Length 735;
  Best Local Similarity   100.0%;  Pred. No. 2.7e-140;
  Matches  735;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         19 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 78
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy         79 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 138
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy        139 ACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 198
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy        199 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 258
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy        259 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 318
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy        319 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 378
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy        379 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 438
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy        439 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 498
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy          499 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 558
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy          559 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 618
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy          619 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy          679 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 738
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy          739 GTCCCCAAAAAGCCG 753
                |||||||||||||||
Db          721 GTCCCCAAAAAGCCG 735


RESULT 13
AAS20787
ID   AAS20787 standard; DNA; 735 BP.
XX
AC   AAS20787;
XX
DT   09-APR-2002  (first entry)
XX
DE   DNA encoding human beta-II tryptase active site mutant H44A #4.
XX
KW   Human; proteolytic tryptase; protease; recombinant beta-II tryptase;
KW   enzyme; mutant; ds.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200198470-A2.
XX
PD   27-DEC-2001.
XX
PF   20-JUN-2001; 2001WO-US019681.
XX
PR   21-JUN-2000; 2000US-00598982.
XX
PA   (PROM-) PROMEGA CORP.
XX
PI   Maffit M,  Niles AL,  Haak-Frendscho M;
XX
DR   WPI; 2002-114578/15.
DR   P-PSDB; AAU12021.
XX
PT   DNA construct for producing enzymatically-inactive proteolytic tryptase,
PT   comprises DNA sequence encoding proteolytic tryptase having an active
PT   site mutation.
XX
PS   Example 1c; Page 115-117; 126pp; English.
XX
CC   The present invention relates to recombinant human proteolytic tryptases,
CC   active site mutants of these tryptases and the methods for producing
CC   these. The method involves the production of a DNA expression construct
CC   comprising a promoter operably linked to a secretion signal sequence
CC   which is operably linked to a DNA sequence encoding a proteolytic
CC   tryptase with an active site mutation (the construct drives expression of
CC   a mature proteolytic tryptase that lacks enzymatic activity due to the
CC   active site mutation, in hosts transformed to contain the construct). The
CC   method is useful for producing enzymatically-active beta-II tryptase. The
CC   active site mutants of proteolytic tryptase provide a tool to investigate
CC   the structural and functional properties of the protease and its
CC   enzymatic activity, and for modelling studies. The enzymatically-active,
CC   recombinant proteolytic tryptase produced are useful as an antigen to
```

```
CC   generate anti-human tryptase antibodies and in drug screening for
CC   compounds which act as tryptase inhibitors, antagonists, agonists, etc.
CC   AAS20775-AAS20790 encode for recombinant human beta-II tryptase active
CC   site mutants
XX
SQ   Sequence 735 BP; 139 A; 246 C; 230 G; 120 T; 0 U; 0 Other;

  Query Match             95.1%;  Score 733.4;  DB 6;  Length 735;
  Best Local Similarity   99.9%;  Pred. No. 5.8e-140;
  Matches  734;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy           19 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 78
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db            1 ATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTC 60

Qy           79 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 138
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           61 CACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTG 120

Qy          139 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 198
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          121 ACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTG 180

Qy          199 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 258
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 CGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCAC 240

Qy          259 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 318
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 CCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCG 300

Qy          319 GTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 378
                ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 GTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCC 360

Qy          379 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 438
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 CCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCA 420

Qy          439 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 498
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 CCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCA 480

Qy          499 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 558
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 AAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTG 540

Qy          559 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 618
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 TGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGC 600

Qy          619 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 678
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 AAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAG 660

Qy          679 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 738
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 CCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTAT 720

Qy          739 GTCCCCAAAAAGCCG 753
                |||||||||||||||
Db          721 GTCCCCAAAAAGCCG 735


RESULT 14
AAV44330
ID   AAV44330 standard; DNA; 1128 BP.
XX
AC   AAV44330;
XX
DT   24-NOV-1998  (first entry)
XX
DE   Human mast cell tryptase II/beta nucleic acid sequence.
```

```
XX
KW   Mast cell tryptase II/beta; human; MCP-7; mast cell protease 7;
KW   blood clot; anticoagulant; myocardial infarction; reocclusion;
KW   thromboembolism; cerebral embolism; thrombosis; therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..826
FT                   /*tag= a
XX
PN   WO9824886-A1.
XX
PD   11-JUN-1998.
XX
PF   25-NOV-1997;   97WO-US021620.
XX
PR   04-DEC-1996;   96US-0032354P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Stevens RL;
XX
DR   WPI; 1998-333308/29.
DR   P-PSDB; AAW64240.
XX
PT   New compositions containing tryptase-7, e.g. mouse mast cell protease-7 -
PT   are used to treat clot formation in e.g. myocardial infarction,
PT   reocclusion following angioplasty or pulmonary thrombo-embolism.
XX
PS   Disclosure; Page 66; 92pp; English.
XX
CC   This nucleotide sequence includes a coding region for human mast cell
CC   tryptase II/beta (see AAW64240). The invention provides: compositions
CC   comprising an isolated tryptase-7 that may include chimeric proteins that
CC   contain (a) a human tryptase for all but the active site region and (b)
CC   the substrate-binding pocket of mouse tryptase-7 or its homologues (see
CC   AAW64233-39); a method for treating a blood clot by administering a
CC   nucleic acid molecule that codes for a tryptase-7, or an expression
CC   product, to decrease fibrinogen activity; a nucleic acid encoding a
CC   serine protease (SP); and a method of producing a mature SP by expressing
CC   the inactive zymogen in a host cell, and cleaving the enterokinase
CC   susceptibility domain. The tryptase-7 polypeptides can be used to treat
CC   disorders mediated by undesirable thrombus clot formation such as
CC   myocardial infarction and reocclusion following angioplasty of blood
CC   clots associated with pulmonary thromboembolism, deep vein thrombosis,
CC   cerebral embolism, renal vein and peripheral arterial thrombosis. They
CC   are also useful for all surgical procedures that require decreased blood
CC   clots
XX
SQ   Sequence 1128 BP; 190 A; 409 C; 329 G; 200 T; 0 U; 0 Other;

  Query Match             94.8%;  Score 731;  DB 2;  Length 1128;
  Best Local Similarity   98.7%;  Pred. No. 1.8e-139;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GCATCGTTGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 146

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 206

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 266

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 326

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 386

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 446

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 506

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 566

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 626

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 686

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 746

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 806

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        807 ATGTCCCCAAAAAGCCGTGAGTCAGGC 833


RESULT 15
AAV42712
ID   AAV42712 standard; cDNA; 1128 BP.
XX
AC   AAV42712;
XX
DT   27-OCT-1998  (first entry)
XX
DE   Human mast cell tryptase II/beta encoding cDNA.
XX
KW   Mast cell protease; MCP; mouse; inhibitor; peptide substrate; asthma;
KW   tryptase-6 protein; inflammatory disorder; allergic rhinitis; urticaria;
KW   antioedema; eczematous dermatitis; atopic dermatitis; anaphylaxis;
KW   hyperproliferative skin disease; peptic ulcer; hyperresponsiveness;
KW   inflammatory skin condition; human; mast cell tryptase II/beta; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             2..826
FT                   /*tag= a
FT                   /product= "Human mast cell tryptase II/beta"
FT                   /note= "the start codon is not indicated"
XX
PN   WO9833812-A1.
XX
PD   06-AUG-1998.
XX
PF   30-JAN-1998;   98WO-US001865.
XX
PR   05-FEB-1997;   97US-0037090P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Stevens RL,  Huang C;
XX
DR   WPI; 1998-437390/37.
```

```
DR   P-PSDB; AAW63175.
XX
PT   Tryptase-6 complex inhibitory peptides - used to treat mast cell-mediated
PT   inflammatory disorders e.g. asthma.
XX
PS   Disclosure; Page 47-48; 69pp; English.
XX
CC   This cDNA encodes the human mast cell tryptase II/beta which is a
CC   homologue of the mouse mast cell protease (mMCP-6) zymogen. The invention
CC   provides sequences shown in AAW63160 to AAW63169 that are inhibitors of
CC   mMCP-6. These peptides which are tryptase-6 complex inhibitors, can be
CC   used for treating a mast cell-mediated inflammatory disorder. The
CC   inhibitors can be used to treat inflammatory disorders including asthma,
CC   allergic rhinitis, urticaria and antioedema, eczematous dermatitis
CC   (atopic dermatitis), anaphylaxis, hyperproliferative skin disease, peptic
CC   ulcers, inflammatory bowel disorder, hyperresponsiveness and inflammatory
CC   skin conditions
XX
SQ   Sequence 1128 BP; 190 A; 409 C; 329 G; 200 T; 0 U; 0 Other;

  Query Match             94.8%;  Score 731;  DB 2;  Length 1128;
  Best Local Similarity   98.7%;  Pred. No. 1.8e-139;
  Matches  737;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         87 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 146

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        147 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 206

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        207 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 266

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        267 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 326

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        327 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 386

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        387 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 446

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        447 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 506

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        507 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 566

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        567 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 626

Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        627 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 686

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 676
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        687 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCC 746

Qy        677 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 736
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        747 AGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCACT 806

Qy        737 ATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              ||||||||||||||||||||  | | |
Db        807 ATGTCCCCAAAAAGCCGTGAGTCAGGC 833

Search completed: July 22, 2004, 11:01:39
Job time : 525 secs
```

```
                    GenCore version 5.1.6
           Copyright (c) 1993 - 2004  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:         July 22, 2004, 10:45:20 ; Search time 3404 Seconds
                                          (without alignments)
                                          6763.731 Million cell updates/sec

Title:          US-09-598-982-20
Perfect score:  771
Sequence:       1 gggcccctcgagaaaagaat..........cgtgaagcggccgccgtcgt 771

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters:       55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  em_gss_hum:*
                 18:  em_gss_inv:*
                 19:  em_gss_pln:*
                 20:  em_gss_vrt:*
                 21:  em_gss_fun:*
                 22:  em_gss_mam:*
                 23:  em_gss_mus:*
                 24:  em_gss_pro:*
                 25:  em_gss_rod:*
                 26:  em_gss_phg:*
                 27:  em_gss_vrl:*
                 28:  gb_gss1:*
                 29:  gb_gss2:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                     SUMMARIES
                 %
Result         Query
  No.   Score  Match Length DB  ID                         Description
 ------------------------------------------------------------------------------
     1   676.2   87.7   1030  12  BM919268                  BM919268 AGENCOURT
     2   581.8   75.5    589  13  BU071733                  BU071733 im33a09.y
     3   580.2   75.3    904  13  BQ720404                  BQ720404 AGENCOURT
     4   574.2   74.5    619  14  CD671851                  CD671851 fg07g11.y
```

```
      5   531.4   68.9    642  14  CD672689                  CD672689 fg15a06.y
      6     525   68.1    635  14  CD671891                  CD671891 fg08b10.y
      7   516.6   67.0    722  13  BX091892                  BX091892 BX091892
      8   495.6   64.3    502  13  BQ082929                  BQ082929 K-EST0144
      9   488.4   63.3    581  13  BU787772                  BU787772 io34c08.y
     10   468.4   60.8    736  12  BG542089                  BG542089 602571309
     11   455.6   59.1    806  14  CF553058                  CF553058 AGENCOURT
     12   444.4   57.6    917  13  BQ721078                  BQ721078 AGENCOURT
     13   430.6   55.8    930  14  CB203717                  CB203717 AGENCOURT
     14   422.2   54.8   1000  12  BI413250                  BI413250 602986390
     15   420.4   54.5    639   9  AA131142                 AA131142 zl31b01.r
c    16     420   54.5    686  13  BU073263                  BU073263 im33a09.x
     17   401.4   52.1    811  12  BG697540                  BG697540 602660760
     18     400   51.9    678  12  BG483591                  BG483591 602503309
     19     375   48.6    512  10  BF724180                  BF724180 bx01h02.y
c    20   371.4   48.2    929  14  CF583026                  CF583026 AGENCOURT
c    21   358.6   46.5    690  14  CA439686                  CA439686 UI-H-DT1-
c    22   357.6   46.4    699  14  CA438721                  CA438721 UI-H-DT1-
c    23     351   45.5    693  12  BM991728                  BM991728 UI-H-DF1-
c    24   350.2   45.4    376  10  BF850308                  BF850308 CM3-EN007
c    25   349.6   45.3    692  12  BM989945                  BM989945 UI-H-DI0-
c    26     348   45.1    688  12  BQ003361                  BQ003361 UI-H-EI1-
c    27     338   43.8    682  14  CA425354                  CA425354 UI-H-DF0-
     28   334.6   43.4    671  14  CB840425                  CB840425 M15E-0715
c    29   334.2   43.3    371  10  BF359130                  BF359130 QV3-ET006
     30   324.2   42.0    859  14  CF583025                  CF583025 AGENCOURT
     31   319.4   41.4    581  14  CA867461                  CA867461 ir28f02.y
c    32   314.6   40.8    379  10  BF849396                  BF849396 CM3-EN007
     33   309.4   40.1    474  10  AW823937                  AW823937 uf61e03.y
     34   305.8   39.7    525  12  BI775716                  BI775716 468400 MA
     35   304.2   39.5    545  12  BI681216                  BI681216 460544 MA
     36     295   38.3    558  12  BI790801                  BI790801 id09c06.y
c    37   289.4   37.5    291  10  BF876081                  BF876081 IL3-ET011
c    38   286.6   37.2    617  14  CA943476                  CA943476 ir28f02.x
     39   281.6   36.5    537  12  BM708916                  BM708916 UI-E-CI1-
c    40   279.6   36.3    620  13  BU731577                  BU731577 UI-E-CI1-
c    41   276.8   35.9    613  10  AW152543                  AW152543 xf76b04.x
     42   273.6   35.5    411  10  BE862270                  BE862270 UI-M-BH0-
     43   266.6   34.6    660  14  CB589106                  CB589106 AGENCOURT
     44   264.8   34.3    538  10  BE751979                  BE751979 204246 MA
     45   261.4   33.9    620   9  AA049080                 AA049080 mj50f09.r


                                  ALIGNMENTS


RESULT 1
BM919268
LOCUS       BM919268                1030 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6715837 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5748499
            5', mRNA sequence.
ACCESSION   BM919268
VERSION     BM919268.1  GI:19369647
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1030)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12776  row: f  column: 20
            High quality sequence stop: 719.
```

```
FEATURES             Location/Qualifiers
     source          1. .1030
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5748499"
                     /lab_host="DH10B"
                     /clone_lib="NIH_MGC_120"
                     /note="Organ: pooled pancreas and spleen; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of spleen and pancreas from 28 yo
                     male.  Library is oligo-dT primed and directionally cloned
                     (EcoRV  site is destroyed upon cloning).  Average insert
                     size 1.5 kb, insert size range 1-2.5 kb.  Library is
                     normalized and enriched for  full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 025. Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             87.7%;  Score 676.2;  DB 12;  Length 1030;
  Best Local Similarity   94.5%;  Pred. No. 1.7e-115;
  Matches  711;  Conservative    0;  Mismatches   39;  Indels    2;  Gaps    1;

Qy          17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GTATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 180

Qy          77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
               ||| || || |||||||||||||||||| |||||||||||||||||||||||||||||||
Db         181 TCCGCGACCGATACTGGATGCACTTCTGTGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 240

Qy         137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
               ||||||| |||   ||| |||||||||||||||||||||||||| |||||||||||||||
Db         241 TGACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAAC 300

Qy         197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 360

Qy         257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
               ||||||||||||||| |  |||||  |||||||| |||||||||||||||||||||||||
Db         361 ACCCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGC 420

Qy         317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
               | |||||  ||||||||| ||||||||||||||  |||||||||||||| ||||||||||
Db         421 CCGTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCC 480

Qy         377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
               |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db         481 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCC 540

Qy         437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
               ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db         541 CACCGCCATTTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 600

Qy         497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
               |||||||||||||||||||||||||||||||||||||||||||| |||||||||||||||
Db         601 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGC 660

Qy         557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
               ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db         661 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCTGGAGGGCCCCTGGTGT 720

Qy         617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGGG--CGAGGGCTGTGC 674
               |||||||||||||||||||||| |||||||||||||||||||||||     |||||||||
Db         721 GCAAGGTGAATGGCACCTGGCTACAGGCGGGCGTGGTCAGCTGGGGACGAGGGGCTGTGC 780

Qy         675 CCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCA 734
               ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         781 CCAGCCCCACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGGACTGGATCCACCA 840

Qy         735 CTATGTCCCCAAAAAGCCGTGAAGCGGCCGCC 766
               |||||||||||||||||||||| |    |||
Db         841 CTATGTCCCCAAAAAGCCGTGAGTCCANGGCC 872


RESULT 2
BU071733
LOCUS       BU071733                 589 bp    mRNA    linear   EST 27-AUG-2002
DEFINITION  im33a09.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6036761
            5' similar to SW:TRYB_HUMAN P20231 BETA-TRYPTASE PRECURSOR ;, mRNA
            sequence.
ACCESSION   BU071733
VERSION     BU071733.1  GI:22512922
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 589)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: im33a09.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 433.
FEATURES             Location/Qualifiers
     source          1. .589
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6036761"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

  Query Match             75.5%;  Score 581.8;  DB 13;  Length 589;
  Best Local Similarity   99.7%;  Pred. No. 4.9e-98;
  Matches  583;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         158 GACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACT 217
               || |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           5 GAGCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCAGCACCTCTACT 64

Qy         218 ACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCC 277
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          65 ACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTTCTACACCGCCC 124
```

```
Qy         278 AGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACG 337
               |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||||
Db         125 AGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACG 184

Qy         338 TCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGG 397
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         185 TCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGG 244

Qy         398 TCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGC 457
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         245 TCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGC 304

Qy         458 AGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCT 517
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         305 AGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCT 364

Qy         518 ACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGA 577
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         365 ACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGA 424

Qy         578 GGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGC 637
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         425 GGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGC 484

Qy         638 TGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCT 697
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         485 TGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCT 544

Qy         698 ACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTATGTCC 742
               |||||||||||||||||||||||||||||||||||||||||||||
Db         545 ACACCCGTGTCACCTACTACTTGGACTGGATCCACCACTATGTCC 589


RESULT 3
BQ720404
LOCUS       BQ720404                 904 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8477973 Lupski_sympathetic_trunk Homo sapiens cDNA clone
            IMAGE:6196776 5', mRNA sequence.
ACCESSION   BQ720404
VERSION     BQ720404.1  GI:21859301
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 904)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13605  row: m  column: 01
            High quality sequence stop: 587.
FEATURES             Location/Qualifiers
     source          1..904
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6196776"
                     /sex="male"
                     /tissue_type="sympathetic trunk"
                     /dev_stage="adult, 16 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_sympathetic_trunk"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
```

```
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.9 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine); available through Life
                     Technologies."
ORIGIN

  Query Match             75.3%;  Score 580.2;  DB 13;  Length 904;
  Best Local Similarity   90.1%;  Pred. No. 1.1e-97;
  Matches  666;  Conservative    0;  Mismatches   43;  Indels   30;  Gaps    3;

Qy          17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
               | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          98 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 157

Qy          77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
               ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         158 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 217

Qy         137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
               ||||||| |||   |||||||||||                           ||||||||
Db         218 TGACCGCAGCGCACTGCGTGGGACC---------------------------GGTGCAAC 250

Qy         197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         251 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 310

Qy         257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         311 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 370

Qy         317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
               ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         371 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 430

Qy         377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         431 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 490

Qy         437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         491 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 550

Qy         497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         551 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 610

Qy         557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         611 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGG 670

Qy         617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGT-GGTCAGCTGGGGCGAGGGCTGTGCC 675
               |||| | ||||||| ||||||||   |||||||| ||||| ||||||||||||||| |||
Db         671 GCAACGCGAATGGCCCCTGGCTGGCAGCGGGCGTGGGTCACCTGGGGCGAGGGCTGGGCC 730

Qy         676 CAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTGG--ACTGGATCCACC 733
               | |||||||||| | ||| |||||||||  | | |  ||  |    |||  |  ||| |||||
Db         731 CGGCCCAACCGGGCGGGCTTCTACACCCCGGGCCCCGTCCTACTGGGAATGGGAGCCACC 790

Qy         734 ACTATGTCCCCAAAAAGCC 752
                || || ||||   ||| |
Db         791 CCTCTGGCCCCCCCAAGAC 809


RESULT 4
CD671851
LOCUS       CD671851                 619 bp    mRNA    linear   EST 24-JUN-2003
DEFINITION  fg07g11.y1 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
            fg07g11 5', mRNA sequence.
```

```
ACCESSION   CD671851
VERSION     CD671851.1  GI:32173582
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 619)
  AUTHORS   Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of adult human iris for the NEIBank
            Project: steroid-response factors and similarities with retinal
            pigment epithelium
  JOURNAL   Mol. Vis. 8 (4), 185-195 (2002)
  MEDLINE   22103462
   PUBMED   12107412
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 07  row: g  column: 11
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..619
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fg07g11"
                     /tissue_type="Iris"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Iris cDNA (Normalized): fg"
                     /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
                     library (bx) was normalized by self-subtraction. One
                     portion of double stranded plasmid DNA representing the
                     library was linearized by NotI. This NotI digested library
                     was used as a template for biotinylated RNA synthesis
                     using SP6 RNA polymerase. Another portion of the double
                     stranded plasmid library was converted to single-stranded
                     circles in vitro using Gene II and Exonuclease III (Life
                     Technologies). Single-stranded DNA (1 mg) was hybridized
                     (C0t 500) with 41 mg of Bio-RNA and vector blocking
                     oligonucleotides. The hybridized Bio-RNA/ss-circles were
                     removed by streptavidin:phenol extraction. EST analysis
                     was performed on the library at the NIH Intramural
                     Sequencing Center(NISC)."
ORIGIN

  Query Match             74.5%;  Score 574.2;  DB 14;  Length 619;
  Best Local Similarity   95.5%;  Pred. No. 1.3e-96;
  Matches  591;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy         26 GGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCACGGCC 85
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| || ||
Db          1 GGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAGTCCGCGACC 60

Qy         86 CATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCCG 145
               |||||||||||||||||| |||||||||||||||||||||||||||||||||||||| |
Db         61 GATACTGGATGCACTTCTGTGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCGG 120

Qy        146 CGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGC 205
              ||   ||| ||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        121 CGCACTGCCTGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAACTGCGGGAGC 180

Qy        206 AGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGT 265
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGT 240

Qy        266 TCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGG 325
              |||||| |  |||||  |||||||| |||||||||||||||||||||||||| |||||
Db        241 TCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACA 300

Qy        326 TCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGA 385
              ||||||||| ||||||||||||||  |||||||||||||| |||||||||||||||||||
Db        301 TCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGA 360

Qy        386 TGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCACCGCCAT 445
              ||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db        361 TGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCACCGCCAT 420

Qy        446 TTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACC 505
              |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 TTCCCCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACC 480

Qy        506 ACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCG 565
              ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db        481 ACCTTGGCGCCTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTGTGTGCCG 540

Qy        566 GGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGA 625
              |||||||||||||||||||||||||||||||||| |||||||||||||||||||||||||
Db        541 GGAACACCCGGAGGGACTCATGCCAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGA 600

Qy        626 ATGGCACCTGGCTGCAGGC 644
              ||||||||||||| |||||
Db        601 ATGGCACCTGGCTACAGGC 619


RESULT 5
CD672689
LOCUS       CD672689                 642 bp    mRNA    linear   EST 24-JUN-2003
DEFINITION  fg15a06.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
            fg15a06 5', mRNA sequence.
ACCESSION   CD672689
VERSION     CD672689.1  GI:32174420
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 642)
  AUTHORS   Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of adult human iris for the NEIBank
            Project: steroid-response factors and similarities with retinal
            pigment epithelium
  JOURNAL   Mol. Vis. 8 (4), 185-195 (2002)
  MEDLINE   22103462
   PUBMED   12107412
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 15  row: a  column: 06
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..642
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fg15a06"
                     /tissue_type="Iris"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Iris cDNA (Normalized): fg"
                     /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
                     library (bx) was normalized by self-subtraction. One
                     portion of double stranded plasmid DNA representing the
                     library was linearized by NotI. This NotI digested library
                     was used as a template for biotinylated RNA synthesis
                     using SP6 RNA polymerase. Another portion of the double
```

```
                    stranded plasmid library was converted to single-stranded
                    circles in vitro using Gene II and Exonuclease III (Life
                    Technologies). Single-stranded DNA (1 mg) was hybridized
                    (C0t 500) with 41 mg of Bio-RNA and vector blocking
                    oligonucleotides. The hybridized Bio-RNA/ss-circles were
                    removed by streptavidin:phenol extraction. EST analysis
                    was performed on the library at the NIH Intramural
                    Sequencing Center(NISC)."
ORIGIN

  Query Match             68.9%;  Score 531.4;  DB 14;  Length 642;
  Best Local Similarity   98.0%;  Pred. No. 1.2e-88;
  Matches  538;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
               | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          94 GCATCGTTGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 153

Qy          77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
               ||| || || ||||||||||||||||||||||||||||||||||||||||||||||||||
Db         154 TCCGCGACCGATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 213

Qy         137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
               ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db         214 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 273

Qy         197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         274 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 333

Qy         257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         334 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 393

Qy         317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
               ||||||| |||||||||||||||||||||||| |||||||||||||||||||||||||||
Db         394 CGGTGAACGTCTCCAGCCACGTCCACACGGTCGCCCTGCCCCCTGCCTCAGAGACCTTCC 453

Qy         377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         454 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 513

Qy         437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         514 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 573

Qy         497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         574 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 633

Qy         557 TGTGTGCCG 565
               |||||||||
Db         634 TGTGTGCCG 642


RESULT 6
CD671891
LOCUS       CD671891                 635 bp    mRNA    linear   EST 24-JUN-2003
DEFINITION  fg08b10.y2 Human Iris cDNA (Normalized): fg Homo sapiens cDNA clone
            fg08b10 5', mRNA sequence.
ACCESSION   CD671891
VERSION     CD671891.1  GI:32173622
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 635)
  AUTHORS   Wistow,G., Bernstein,S.L., Ray,S., Wyatt,M.K., Behal,A.,
            Touchman,J.W., Bouffard,G., Smith,D. and Peterson,K.
  TITLE     Expressed sequence tag analysis of adult human iris for the NEIBank
            Project: steroid-response factors and similarities with retinal
            pigment epithelium
  JOURNAL   Mol. Vis. 8 (4), 185-195 (2002)
   MEDLINE  22103462
    PUBMED  12107412
COMMENT     Contact: Wistow G
            Section on Molecular Structure and Function
            National Eye Institute
            6/331, NIH, Bethesda, MD 20892-2740, USA
            Tel: 301 402 3452
            Fax: 301 496 0078
            Email: graeme@helix.nih.gov
            Plate: 08  row: b  column: 10
            Seq primer: M13RP1 reverse primer (ABI).
FEATURES             Location/Qualifiers
     source          1..635
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="fg08b10"
                     /tissue_type="Iris"
                     /dev_stage="Adult"
                     /lab_host="EMDH10B"
                     /clone_lib="Human Iris cDNA (Normalized): fg"
                     /note="Organ: Eye; Vector: pCMVSPORT6; A human iris
                     library (bx) was normalized by self-subtraction. One
                     portion of double stranded plasmid DNA representing the
                     library was linearized by NotI. This NotI digested library
                     was used as a template for biotinylated RNA synthesis
                     using SP6 RNA polymerase. Another portion of the double
                     stranded plasmid library was converted to single-stranded
                     circles in vitro using Gene II and Exonuclease III (Life
                     Technologies). Single-stranded DNA (1 mg) was hybridized
                     (C0t 500) with 41 mg of Bio-RNA and vector blocking
                     oligonucleotides. The hybridized Bio-RNA/ss-circles were
                     removed by streptavidin:phenol extraction. EST analysis
                     was performed on the library at the NIH Intramural
                     Sequencing Center(NISC)."
ORIGIN

  Query Match             68.1%;  Score 525;  DB 14;  Length 635;
  Best Local Similarity   98.2%;  Pred. No. 1.8e-87;
  Matches  531;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
               | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          95 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 154

Qy          77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
               |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         155 TCCATGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 214

Qy         137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
               ||  ||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db         215 TGCTCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 274

Qy         197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         275 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 334

Qy         257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
               |||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         335 ACCCCCAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 394

Qy         317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
               ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         395 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 454

Qy         377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         455 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 514

Qy         437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         515 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 574
```

```
Qy         497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         575 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 634

Qy         557 T 557
               |
Db         635 T 635

RESULT 7
BX091892
LOCUS       BX091892                 722 bp    mRNA    linear   EST 23-JAN-2003
DEFINITION  BX091892 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGp998M0584 ; IMAGE:110380, mRNA sequence.
ACCESSION   BX091892
VERSION     BX091892.1  GI:27822632
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 722)
  AUTHORS   Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Human UnigeneSet - RZPD3
  JOURNAL   Unpublished (2003)
COMMENT     Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
            RZPD; IMAGp998M0584.
            RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
            Human UnigeneSet - RZPD3 (RZPDLIB No.972)
            http://www.rzpd.de/CloneCards/cgi-
            bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 101
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            M13r, Primer sequence: TTTCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .722
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGp998M0584 ; IMAGE:110380"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
ORIGIN

  Query Match            67.0%;  Score 516.6;  DB 13;  Length 722;
  Best Local Similarity   99.2%;  Pred. No. 6.8e-86;
  Matches  519;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         241 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 AGCAGGATCATCGTGCACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTG 60

Qy         301 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 CTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCT 120

Qy         361 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 GCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGAC 180

Qy         421 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         181 AATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAA 240

Qy         481 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         241 AACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATC 300

Qy         541 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         301 GTCCGTGACGACATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCC 360

Qy         601 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         361 GGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGG 420

Qy         661 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         421 GGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTACTTG 480

Qy         721 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCC 763
               ||||||||||||||||||||||||||||||||||||  | | |
Db         481 GACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAGTCAGGC 523

RESULT 8
BQ082929
LOCUS       BQ082929                 502 bp    mRNA    linear   EST 04-APR-2002
DEFINITION  K-EST0144841 S14K402 Homo sapiens cDNA clone S14K402-59-B11 5',
            mRNA sequence.
ACCESSION   BQ082929
VERSION     BQ082929.1  GI:19941018
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 502)
  AUTHORS   Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
            Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
            Kim,Y.S.
  TITLE     21C Frontier Korean EST Project 2001
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Kim YS
            Genome Research Center
            Korea Research Institute of Bioscience & Biotechnology
            52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
            Tel: +82-42-860-4470
            Fax: +82-42-860-4409
            Email: yongsung@mail.kribb.re.kr
            Plate: 59  row: B  column: 11
            High quality sequence stop: 502.
FEATURES             Location/Qualifiers
     source          1. .502
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="S14K402-59-B11"
                     /cell_line="K402"
                     /lab_host="Top10F'"
                     /clone_lib="S14K402"
                     /note="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
                     Site_2: NotI; The poly (A)+ RNA was dephosphorylated with
                     bacterial alkaline phosphatase (BAP) and then decapped
                     with tobacco acid pyrophosphatase (TAP). The decapped
                     intact mRNA was ligated with DNA-RNA linker including EcoR
```

```
                I site by treatment of T4 RNA ligase and the first strand
                cDNA was synthesized from oligo dT-selected mRNA by
                priming with dT-tailed vector. The dT-tailed vector was
                adjusted to have about 60nt. The cDNA vector was
                circularized with E. coli DNA ligase after digestion of
                EcoRI which site is also included in vector. An RNA strand
                converted to a DNA strand by Okayama-Berg method. The
                obtained cDNA vectors were used for transformation of
                competent cells E. coli Top10F' by electroporation method.
                The cDNA libraries constructed by this method are
                full-length enriched cDNA library."
ORIGIN

  Query Match             64.3%;  Score 495.6;  DB 13;  Length 502;
  Best Local Similarity   99.2%;  Pred. No. 4.7e-82;
  Matches  498;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         87 ATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCCGC 146
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db          1 ATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGCTGACCGCAGC 60

Qy        147 GGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCA 206
              |   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAACTGCGGGAGCA 120

Qy        207 GCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTT 266
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 GCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGCACCCACAGTT 180

Qy        267 CTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGT 326
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 CTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGT 240

Qy        327 CTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGAT 386
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 CTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGAT 300

Qy        387 GCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATT 446
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 GCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATT 360

Qy        447 TCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCA 506
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCA 420

Qy        507 CCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGG 566
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 CCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGG 480

Qy        567 GAACACCCGGAGGGACTCATGC 588
              ||||||||||||||||||||||
Db        481 GAACACCCGGAGGGACTCATGC 502


RESULT 9
BU787772
LOCUS       BU787772                 581 bp    mRNA    linear   EST 11-OCT-2002
DEFINITION  io34c08.y1 Human insulinoma Homo sapiens cDNA clone IMAGE:6128247
            5' similar to SW:TRYB_HUMAN P20231 BETA-TRYPTASE PRECURSOR ;, mRNA
            sequence.
ACCESSION   BU787772
VERSION     BU787772.1  GI:23835713
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 581)
  AUTHORS   Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
            Lemishka,I., Scearce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
            Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
            Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
            Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagareishvili,R.,
```

```
            Williams,T., Jackson,Y. and Bowers,Y.
  TITLE     Endocrine Pancreas Consortium
  JOURNAL   Unpublished (2000)
COMMENT     Other_ESTs: io34c08.x1
            Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
            Endocrine Pancreas Consortium
            Harvard University, Howard Hughes Medical Institute
            Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
            MA 02138
            Tel: 617-495-1812
            Fax: 617-495-8557
            Email: dmelton@biohp.harvard.edu
            Library was constructed by Dr. J. Ferrer In vivo mass-excised to
            pBluescript SK- by Dr. H. Inoue DNA sequencing by: Washington
            University Genome Sequencing Center For information on obtaining a
            clone please contact: Dr. Hiroshi Inoue (hinoue@im.wustl.edu)
            Seq primer: -40RP from Gibco
            High quality sequence stop: 427.
FEATURES             Location/Qualifiers
     source          1. .581
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6128247"
                     /tissue_type="insulinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="Human insulinoma"
                     /note="Organ: pancreas; Vector: pBluescript SK-; Site_1:
                     XhoI; Site_2: EcoRI; Constructed with lambda ZAPII system
                     (Stratagene) by Dr. J. Ferrer, in vivo mass-excised to
                     pBluescript SK- by Dr. H. Inoue  following the Washington
                     University protocol
                     (http://genome.wustl.edu/est/lambda_protocol.shtml).
                     Please contact  Hiroshi Inoue, MD/PhD for further
                     information on this library  (Metabolism Division, Permutt
                     Laboratory, Washington University School  of Medicine, Box
                     8127, 660 S Euclid Ave, St. Louis, MO 63110).  Note: this
                     is a Washington University Pancreas EST project library. "
ORIGIN

  Query Match             63.3%;  Score 488.4;  DB 13;  Length 581;
  Best Local Similarity   98.8%;  Pred. No. 1.1e-80;
  Matches  492;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         84 GCATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 143

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        144 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 203

Qy        137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||||||||||||||||||||||||||||||||||||||||||||||
Db        204 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 263

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        264 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 323

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        324 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 383

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        384 CGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 443

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        444 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 503

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         504 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 563

Qy         497 CAAAATACCACCTTGGCG 514
               ||||||||||||||||||
Db         564 CAAAATACCACCTTGGCG 581


RESULT 10
BG542089
LOCUS       BG542089                 736 bp    mRNA    linear   EST 03-APR-2001
DEFINITION  602571309F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4695458 5',
            mRNA sequence.
ACCESSION   BG542089
VERSION     BG542089.1  GI:13534322
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 736)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
             cDNA Library Preparation: CLONTECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1521  row: b  column: 03
            High quality sequence stop: 725.
FEATURES             Location/Qualifiers
     source          1..736
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4695458"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /clone_lib="NIH_MGC_77"
                     /note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc); 5' and
                     3' adaptors were used in cloning as follows: 5' adaptor
                     sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:
                     5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3' (where B = A,
                     C, or G and N = A, C,  G, or T).  Average insert size 1.9
                     kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
                     by PCR. This library was enriched for full-length clones
                     and was constructed by Clontech  Laboratories (Palo Alto,
                     CA). Note: this is a NIH_MGC Library."
ORIGIN

  Query Match             60.8%;  Score 468.4;  DB 12;  Length 736;
  Best Local Similarity   90.9%;  Pred. No. 6.1e-77;
  Matches  567;  Conservative    0;  Mismatches   46;  Indels   11;  Gaps    6;

Qy          17 GAATCGTCGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
               | ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         113 GCATCGTTGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 172

Qy          77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         173 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 232

Qy         137 TGACCGCCGCGGCGTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
               ||||||| |||    |||||||||||||||||||||||||||||||||||||||||||||
Db         233 TGACCGCAGCGCACTGCGTGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 292

Qy         197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         293 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 352

Qy         257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         353 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 412

Qy         317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         413 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 472

Qy         377 CCCC-GGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTC 435
               |||| |||||||||||||||||||||||||||||||||||||||| |          |
Db         473 CCCCGGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACCAATGATGAGCGCCT 532

Qy         436 CCACCGCCATTTCCTCTGAAGCAGG-TGAAGGTCCCCATAATGGAAAACCACATTTGTGA 494
               |||| ||   || | |||||||||| ||||||||||||||||||||||||||||||||||
Db         533 CCACGGCATTTCCTTCTGAAGCAGGTTGAAGGTCCCCATAATGGAAAACCACATTTGTGA 592

Qy         495 CGCAAAATACCACCTTGGCGCCT-ACACGGAGACG-ACGTCCGCATCGTCCGTGACGAC 552
               ||||||||| |||||||| |||| |||||||||||| | ||||||||     |  |  ||
Db         593 CGCAAAATAACACCTTGGGGCCTAACACGGAGACGAAGGTCCGCATTCGTCCGGTGAAC 652

Qy         553 ATGCT-----GTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCG--ACTCCGGAGG 605
                   |      |||||||||||||||||||||||||||||||||||||     | |||||
Db         653 GACATTGCTGTTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGGACTCCGGGAGG 712

Qy         606 GCCCCTGGTGTGCAAGGTGAATGG 629
               |||||||||||||||||||| |||
Db         713 GCCCCTGGTGTGCAAGGTGACTGG 736


RESULT 11
CF553058
LOCUS       CF553058                 806 bp    mRNA    linear   EST 22-SEP-2003
DEFINITION  AGENCOURT_15594982 NIH_MGC_183 Homo sapiens cDNA clone
            IMAGE:30529468 5', mRNA sequence.
ACCESSION   CF553058
VERSION     CF553058.1  GI:34889892
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 806)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Daniela S. Gerhard, Ph.D.
            Office of Cancer Genomics
            National Cancer Institute / NIH
            Bldg. 31 Rm10A07 Bethesda, MD 20892
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Michael Brownstein
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM617  row: b  column: 05
            High quality sequence start: 18
            High quality sequence stop: 611.
FEATURES             Location/Qualifiers
     source          1..806
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30529468"
                     /lab_host="DH10B-TonA (T1 and T5 phage resistant)"
                     /clone_lib="NIH_MGC_183"
                     /note="Organ: Pooled muscle (cardiac and skeletal);
                     Vector: pCMV-SPORT6.1; Site_1: EcoRV (destroyed); Site_2:
                     NotI; Library is oligo-dT primed and directionally cloned
```

```
                     (EcoRV site is destroyed upon cloning). Average insert
                     size 1.7. Library was constructed by Invitrogen."
ORIGIN

  Query Match             59.1%;  Score 455.6;  DB 14;  Length 806;
  Best Local Similarity   94.0%;  Pred. No. 1.5e-74;
  Matches  485;  Conservative    0;  Mismatches   29;  Indels    2;  Gaps    1;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        246 GTATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 305

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              ||| || || |||||||||||||||||| |||||||||||||||||||||||||||||||
Db        306 TCCGCGACCGATACTGGATGCACTTCTGTGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 365

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              ||||||| |||   ||| |||||||||||||||||||||||||| |||||||||||||||
Db        366 TGACCGCGGCGCACTGCCTGGGACCGGACGTCAAGGATCTGGCCACCCTCAGGGTGCAAC 425

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        426 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 485

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              ||||||||||||||| |  |||||  |||||||| |||||||||||||||||||||||||
Db        486 ACCCACAGTTCTACATCATCCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGC 545

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | |||||  ||||||||| ||||||||||||||  |||||||||||||| ||||||||||
Db        546 CCGTGAACATCTCCAGCCGCGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCC 605

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db        606 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCC 665

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              ||||||||||||| |||||||||||||| |||||||||||||||||||||||||||||||
Db        666 CACCGCCATTTCCCCTGAAGCAGGTGAAAGTCCCCATAATGGAAAACCACATTTGTGACG 725

Qy        497 CAAAATACCACCTTGG--CGCCTACACGGGAGACGA 530
              ||||||||||||||||    ||||||||||||||||
Db        726 CAAAATACCACCTTGGGCGCCCTACACGGGAGACGA 761


RESULT 12
BQ721078
LOCUS       BQ721078                 917 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8109065 Lupski_sympathetic_trunk Homo sapiens cDNA clone
            IMAGE:6189508 5', mRNA sequence.
ACCESSION   BQ721078
VERSION     BQ721078.1  GI:21859975
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 917)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13586  row: n  column: 05
            High quality sequence stop: 609.
FEATURES             Location/Qualifiers
```

```
     source          1..917
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6189508"
                     /sex="male"
                     /tissue_type="sympathetic trunk"
                     /dev_stage="adult, 16 yr"
                     /lab_host="DH10B"
                     /clone_lib="Lupski_sympathetic_trunk"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.9 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated  by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine); available through Life
                     Technologies."
ORIGIN

  Query Match             57.6%;  Score 444.4;  DB 13;  Length 917;
  Best Local Similarity   98.5%;  Pred. No. 1.9e-72;
  Matches  459;  Conservative    0;  Mismatches    6;  Indels    1;  Gaps    1;

Qy        299 TGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCC 358
              ||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db          1 TGCTGGAGCTGGAGGAGCCGGTGAACGTCTCCAGCCACGTCCACACGGTCACCCTGCCCC 60

Qy        359 CTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGG 418
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 CTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGG 120

Qy        419 ACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGG 478
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGG 180

Qy        479 AAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCA 538
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 AAAACCACATTTGTGACGCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCA 240

Qy        539 TCGTCCGTGACGA-CATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGAC 597
              ||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db        241 TCGTCCGTGACGAGCATGCTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGAC 300

Qy        598 TCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGC 657
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 TCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGC 360

Qy        658 TGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTAC 717
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCATCTACACCCGTGTCACCTACTAC 420

Qy        718 TTGGACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAAGCGGCC 763
              | |||||||||||||||||||||||||||||||||||||  | | |
Db        421 TGGGACTGGATCCACCACTATGTCCCCAAAAAGCCGTGAGTCAGGC 466


RESULT 13
CB203717
LOCUS       CB203717                 930 bp    mRNA    linear   EST 05-FEB-2003
DEFINITION  AGENCOURT_11288628 NIH_MGC_135 Mus musculus cDNA clone
            IMAGE:30142408 5', mRNA sequence.
ACCESSION   CB203717
VERSION     CB203717.1  GI:28240351
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
```

```
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. David Rowe
             cDNA Library Preparation: Invitrogen Corp
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM0051  row: j  column: 17
            High quality sequence stop: 681.
FEATURES             Location/Qualifiers
     source          1..930
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:30142408"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_135"
                     /note="Vector: pCMVSport6.1; Site_1: EcoRV; Site_2: NotI;
                     Normalized full-length enriched library from pooled mouse
                     embryonic limb, maxilla and mandible, day 12.5, 13.5,
                     14.5, and 15.5 (size selected for the 0.5-1 kb fragments)
                     Cloned directionally, priming method: Oligo-dT. cDNA
                     enrichment: >1k bp, Average insert size 1.6k bp.
                     Normalization (Cot value): 7.5 kb. Priming sequence:
                     5'GACTAGTTCTAGATCGCGAGCGGCCGCCC(T)3' Tissue contributed
                     by, David Rowe. Library constructed by ResGen, Invitrogen
                     Corp."
ORIGIN

  Query Match             55.8%;  Score 430.6;  DB 14;  Length 930;
  Best Local Similarity   76.7%;  Pred. No. 7e-70;
  Matches  552;  Conservative    0;  Mismatches  165;  Indels    3;  Gaps    2;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| || || || |||||  |   ||| |||||||||||||||||||||||||||
Db         40 GCATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGAT 99

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
              |  |       |||||||| || |||||||| ||||| ||||||||||| ||||||||||
Db        100 TTAAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGC 159

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              | || || ||    || ||||||||| || ||||    |       | || ||||||| |
Db        160 TCACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGC 219

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              | || |||||| | || |||||   ||||||||| ||  |  | | | ||||| | ||||
Db        220 TTCGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGC 279

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              |||| || |  ||||| ||| ||   || || |   |  ||||||||||||| ||||  |
Db        280 ACCCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCC 339

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | ||||| ||||||| |||  ||||| |  |  |||||||||||||||| ||||||||||
Db        340 CTGTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCC 399

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              |||| ||||   |||||||||| || |||||||||||  | || ||||| ||||  | ||
Db        400 CCCCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCC 459

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACG 496
              |||| || | |||||||||||| |||||||| |||||  ||||||||  | | ||||||
Db        460 CACCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACC 519

Qy        497 CAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATGC 556
                || ||||||  ||||  |||||||||||| ||  ||| ||| |||| ||| | |||||
Db        520 GGAAGTACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATGC 579
```

```
Qy        557 TGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGT 616
              ||||||| || || ||| |||| ||||| ||||||||||| || || ||||| ||||| |
Db        580 TGTGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTCT 639

Qy        617 GCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCAGCTGGG-GCGAGGGCTGTGCC 675
              |||| ||||| || |||||||||||||| || ||||||||||||| | |||||||| ||
Db        640 GCAAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCAGCTGGGNGTGAGGGCTGCGCA 699

Qy        676 CAGCCCAACCGGCCTGGCATCTACACCCGTGTCACC--TACTACTTGGACTGGATCCACC 733
              |||||||||  |||||||||||||||||| |    |    |||||| |||||||||||||
Db        700 CAGCCCAACAAGCCTGGCATCTACACCCGGGGTGACATANCTACTTAGACTGGATCCACC 759

RESULT 14
BI413250
LOCUS       BI413250                1000 bp    mRNA    linear   EST 14-AUG-2001
DEFINITION  602986390F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5142478 5',
            mRNA sequence.
ACCESSION   BI413250
VERSION     BI413250.1  GI:15174173
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1000)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by:Incyte Genomics, Inc.
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11351  row: c  column: 23
            High quality sequence start: 32
            High quality sequence stop: 849.
FEATURES             Location/Qualifiers
     source          1..1000
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="Czech II"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:5142478"
                     /tissue_type="pooled lung tumors"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NCI_CGAP_Lu33"
                     /note="Organ: lung; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
                     strand cDNA was prepared from mRNA obtained from pooled
                     lung tumors with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCCTCTGTTTTTTTTTTTTTTTTT 3'].
                     Double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia),  digested with Not I and cloned into the Not
                     I and Eco RI sites of  the modified pT7T3 vector. Library
                     went through one round of normalization, and was
                     constructed  by Bento Soares and M. Fatima Bonaldo. "
ORIGIN

  Query Match             54.8%;  Score 422.2;  DB 12;  Length 1000;
  Best Local Similarity   77.2%;  Pred. No. 2.6e-68;
  Matches  564;  Conservative    0;  Mismatches  163;  Indels    4;  Gaps    4;

Qy         17 GAATCGTCGGGGGTCAGGAGGCCCCCAGGAGCAAGTGGCCCTGGCAGGTGAGCCTGAGAG 76
              | ||||| || || || |||||  |   ||| |||||||||||||||||||||||||||
Db         73 GCATCGTGGGAGGACATGAGGCTTCTGAGAGTAAGTGGCCCTGGCAGGTGAGCCTGAGAT 132

Qy         77 TCCACGGCCCATACTGGATGCACTTCTGCGGGGGCTCCCTCATCCACCCCCAGTGGGTGC 136
```

```
              |  |       |||||||| || |||||||| ||||| ||||||||||| ||||||||||
Db        133 TTAAATTAAACTACTGGATACATTTCTGCGGAGGCTCTCTCATCCACCCACAGTGGGTGC 192

Qy        137 TGACCGCCGCGGCGTGCGTGGGACCGGACGTCAAGGATCTGGCCGCCCTCAGGGTGCAAC 196
              | || || ||    || ||||||||| || ||||    |       | || ||||||| |
Db        193 TCACTGCGGCACACTGTGTGGGACCGCACATCAAAAGCCCACAGCTCTTCCGGGTGCAGC 252

Qy        197 TGCGGGAGCAGCACCTCTACTACCAGGACCAGCTGCTGCCGGTCAGCAGGATCATCGTGC 256
              | || ||||||| | || |||||   ||||||||| ||  |  | | | ||||| | ||||
Db        253 TTCGTGAGCAGTATCTATACTATGGGGACCAGCTCCTCTCTTTGAACCGGATCGTGGTGC 312

Qy        257 ACCCACAGTTCTACACCGCCCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGC 316
              |||| || |  ||||| ||| ||   || || ||| | |||||||||||||| ||||  |
Db        313 ACCCCCACTATTACACGGCCGAGGGTGGGGCAGACGTTGCCCTGCTGGAGCTTGAGGTCC 372

Qy        317 CGGTGAAGGTCTCCAGCCACGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCC 376
              | ||||| ||||||| |||  ||||| |  |  |||||||||||||||| ||||||||||
Db        373 CTGTGAATGTCTCCACCCATATCCACCCCATATCCCTGCCCCCTGCCTCGGAGACCTTCC 432

Qy        377 CCCCGGGGATGCCGTGCTGGGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCC 436
              |||| ||||   |||||||||| || |||||||||||  | || ||||| ||||  ||||
Db        433 CCCCTGGGACATCGTGCTGGGTGACAGGCTGGGGCGACATTGATAATGACGAGCCTCTCC 492

Qy        437 CACCGCCATTTCCTCTGAAGCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGA-C 495
              |||| || | |||||||||||| |||||||| |||||  ||||||||  | | ||||| |
Db        493 CACCTCCTTATCCTCTGAAGCAAGTGAAGGTTCCCATTGTGGAAAACAGCCTGTGTGACC 552

Qy        496 GCAAAATACCACCTTGGCGCCTACACGGGAGACGACGTCCGCATCGTCCGTGACGACATG 555
              | ||   |||||  ||||  |||||||||||| ||  | | ||| |||| ||| | ||||
Db        553 GGAAGTAACCACACTGGCCTCTACACGGGAGATGATTTTCCCATTGTCCATGATGGCATG 612

Qy        556 CTGTGTGCCGGGAACACCCGGAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTG 615
              |||||||| ||| || ||| |||| ||||| ||||||||||| || || ||||| |||||
Db        613 CTGTGTGCTGGAAATACCAGGAGAGACTCCTGCCAGGGCGATTCAGGGGGGCCACTGGTC 672

Qy        616 TGCAAGGTGAATGGCACCTGGCTGCAGGCGGGCGTGGTCA-GCTGGGGCGAGGGCTGTGC 674
              ||||| ||||| || |||||||||||||| || ||||||| ||||||| |||||||| ||
Db        673 TGCAAAGTGAAGGGTACCTGGCTGCAGGCAGGAGTGGTCACGCTGGGGTGAGGGCTGCGC 732

Qy        675 CCAGCCC-AACCGGCCTGGCATCTACACCCGTGTCACC-TACTACTTGGACTGGATCCAC 732
               |||||| |||  |||||||||||||||||| || ||| |||||||| ||||||||||||
Db        733 ACAGCCCAAACAAGCCTGGCATCTACACCCGGGTGACCATACTACTTAGACTGGATCCAC 792

Qy        733 CACTATGTCCC 743
                 ||||||||
Db        793 GCTTATGTCCC 803


RESULT 15
AA131142
LOCUS       AA131142                  639 bp    mRNA    linear   EST 14-MAY-1997
DEFINITION  zl31b01.r1 Soares_pregnant_uterus_NbHPU Homo sapiens cDNA clone
            IMAGE:503497 5' similar to gb:M30038 ALPHA-TRYPTASE PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION   AA131142
VERSION     AA131142.1  GI:1692796
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 639)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W.,
            Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,
            Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L.,
            Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J.,
            Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R.
            and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
   PUBMED   8889549
```

```
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 894   Std Error: 0.00
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 349.
FEATURES             Location/Qualifiers
     source          1. .639
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="GDB:3808485"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:503497"
                     /sex="female"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /clone_lib="Soares_pregnant_uterus_NbHPU"
                     /note="Organ: uterus; Vector: pT7T3-Pac; Site_1: Not I;
                     Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCCTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization.  Library
                     constructed by M. Fatima Bonaldo."
ORIGIN

  Query Match             54.5%;  Score 420.4;  DB 9;  Length 639;
  Best Local Similarity   94.0%;  Pred. No. 4.8e-68;
  Matches  458;  Conservative    0;  Mismatches   27;  Indels    2;  Gaps    2;

Qy        276 CCAGATCGGAGCGGACATCGCCCTGCTGGAGCTGGAGGAGCCGGTGAAGGTCTCCAGCCA 335
              |||||  |||||||| |||||||||||||||||||||||||| |||||  |||||||||
Db          1 CCAGACTGGAGCGGATATCGCCCTGCTGGAGCTGGAGGAGCCCGTGAACATCTCCAGCCG 60

Qy        336 CGTCCACACGGTCACCCTGCCCCCTGCCTCAGAGACCTTCCCCCCGGGGATGCCGTGCTG 395
              ||||||||||||||  |||||||||||||| |||||||||||||||||||||||||||||
Db         61 CGTCCACACGGTCATGCTGCCCCCTGCCTCGGAGACCTTCCCCCCGGGGATGCCGTGCTG 120

Qy        396 GGTCACTGGCTGGGGCGATGTGGACAATGATGAGCGCCTCCCACCGCCATTTCCTCTGAA 455
              ||||||||||||||||||||||||||||||||||| |||||||||||||||||| |||||
Db        121 GGTCACTGGCTGGGGCGATGTGGACAATGATGAGCCCCTCCCACCGCCATTTCCCCTGAA 180

Qy        456 GCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGC 515
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 GCAGGTGAAGGTCCCCATAATGGAAAACCACATTTGTGACGCAAAATACCACCTTGGCGC 240

Qy        516 CTACACGGAGACGACGTCCGCATCGTCCGTGACGACATGCTGTGTGCCGGGAACACCCG 575
              |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        241 CTACACGGGAGACGACGTCCGCATCATCCGTGACGACATGCTGTGTGCCGGGAACACCCG 300

Qy        576 GAGGGACTCATGCCAGGGCGACTCCGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTG 635
              ||| |  ||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db        301 GAGAGCTTCATGCCAGGGCGACTCTGGAGGGCCCCTGGTGTGCAAGGTGAATGGCACCTG 360

Qy        636 GCTGCAGGCGGGCGTGGTCAGCTGGGGCGAGGGCTGTGCCCAGCCCAACCGGCCTGGCAT 695
              ||| |||||||||||||||| ||||| ||||||||||| |   |||||||||||||||||
Db        361 GCTACAGGCGGGCGTGGTCATCTGGGACGAGGGCTGTG-CAGTCCCAACCGGCCTGGCAT 419

Qy        696 CTACACCCGTGTCACCTACTACTTGG-ACTGGATCCACCACTATGTCCCCAAAAAGCCGT 754
              |||||||||||||||||||||||||| ||||||||||||| |||||||||||||||||||
Db        420 CTACACCCGTGTCACCTACTACTTGGAACTGGATCCACCANTATGTCCCCAAAAAGCCGT 479

Qy        755 GAAGCGG 761
              ||  | |
Db        480 GAGTCAG 486
```

```
Search completed: July 22, 2004, 13:24:27
Job time : 3410 secs
```